Protein Structure and Dynamics probed by Multiparameter Fluorescence

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Julian Folz

Düsseldorf, August 2023

Institute of Molecular Physical Chemistry
Heinrich-Heine-University Düsseldorf

Printed with the permission of
Department of Mathematics and Natural Sciences
Heinrich-Heine-University Düsseldorf

1. Referee: Prof. Dr. Claus A.M. Seidel
2. Referee: Prof. Dr. Christian Herrmann
3. Referee: Prof. Dr. Jörg Fitter

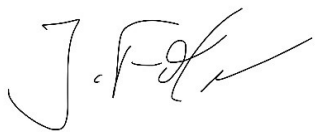
Date of oral examination: 08.08.2023

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Diese Dissertation wurde in der vorgelegten oder einer ähnlichen Form noch bei keiner anderen Institution eingereicht, und es wurden bisher keine erfolglosen Promotionsversuche von mir unternommen.

Düsseldorf, 22.08.2023



Abstract

The structure and dynamical behavior of biomolecules are key to the understanding of their function and underlying mechanisms. An important experimental method to study structural dynamics over 12 orders of magnitude in time is single-molecule (sm) fluorescence spectroscopy.

The use of more than one fluorophore per molecule opens additional opportunities arising from photon densities, coincidences and dipolar coupling by Förster Resonance Energy Transfer (FRET) to study the stoichiometry, structure, dynamics and conformational transitions of biomolecules. One advantage of the method is that molecules can be studied under ambient conditions using single-molecule concentrations in the pM range. Hence it is possible to follow the mechanism of for instance a single protein that it performs in order to fulfil their biological function.

The following thesis is aimed to contribute to discussed topics addressing two different areas, a methodological one and a following application-based one answering open questions related to specific biomolecules. In the first main chapter the accuracy of the method was benchmarked in a worldwide smFRET study on proteins. In the second main chapter I applied smFRET measurements to study various biomolecules performing a wide range of biological functions involved in genetic recombination, the human immune response and how a toxin perforates the membrane of a cell.

The results of the worldwide smFRET study yielded good agreement between the different labs when studying the structure and dynamics of biomolecules. In the study two proteins were used with two specific purposes. The first one was used to study the performance of measurement and analysis methods to derive distances from the FRET measurements. The second one was used as a model system for the description of complex conformational dynamics. However, in this study, user bias and non-uniform calibration was observed. In order to resolve these issues, a follow-up study presents a workflow for more robust calibration. Finally, a new idea is introduced which can potentially reduce the effect of dye artifacts occurring from undesired interaction with the surface of the biomolecule it is attached to.

For the application of smFRET on biomolecules, I studied various biomolecules that are influenced by different environmental conditions. For the Holliday Junction I was able to track the conformation changes it performs on a sub-ms timescale when adding magnesium ions. To this end, a novel device was used enhancing the obtained fluorescence signal. As a second biologically relevant system, I investigated the influence of a small farnesyl moiety on the human guanylate binding protein 1. For this protein, deeper insights in the pathway from a monomeric to its oligomeric state were gained. The last biomolecule investigated was a large protein complex, the Tc toxin, for which new insights into the syringe-like mechanism were resolved. Here, I combined findings from a range of experimental fluorescence methods and was able to derive a consistent picture of the functional mechanism that involved different stable and transient intermediate states.

The obtained results of the global FRET study and the subsequently developed calibration workflow will aid in unifying measurement and analysis methods within the smFRET community, thereby allowing for more consistent and trustworthy smFRET studies on biomolecules. The presented investigation of biomolecules will help to understand basic steps of a fundamental system in genetics, show a protein performing a whole structural pathway to follow its function in the immune response, and how perforating a cell in a syringe like mechanism is performed by a toxin, that has possibly applications as a biopesticide.

Acknowledgments

First of all, I want to thank Prof. Dr. Claus A.M. Seidel, an outstanding and passionate scientist, who offered me to work on world class projects on which I could apply my high frustration tolerance I gained during my studies, leading to exciting science. Secondly, I want to thank my mentor Prof. Dr. Stefan U. Egelhaaf not only mentoring/supervising this thesis but also all previous ones during my academic career, including guidance whenever I needed it.

This thesis is based on many different collaborative projects, for which I would like to thank my main collaborators: For chapter A I want to thank Milana Popara for proof reading and her excellent contributions to our common projects. For chapter B I want to thank L. Morales-Inostroza, Prof. S. Götzinger and Prof. V. Sandoghdar for working with us on new developments and their trust that we will make the most out of it. I want to thank Dr. P. Lauterjung and Prof. C. Herrmann for our common hGBP1 project which is based on their high-quality samples. Furthermore, I want to thank Dr. D. Roderer, Dr. P. Njenga Ng'ang'a and Prof. S. Raunser for a fruitful collaboration investigating the exciting and complex Tc toxin.

Next, I want to thank all people who proofread this thesis, which are Alexander Larbig, Milana Popara, Dr. Christian Hanke, Dr. Ralf Kühnemuth, Luis Morales-Inostroza, Dr. Daniel Roderer, Dr. Paul Lauterjung, Dr. Thorsten Auth, Tim Vogel, Florian Königs and Caroline Kerscht.

Furthermore, I want to thank the whole AG Seidel group and especially Dr. Suren Felekyan, Dr. Ralf Kühnemuth and Dr. Thomas Peulen for their support and an excellent working atmosphere.

Finally, I want to thank my family on which I can always count on and Caroline for her support during my studies and the joy she brings into my life.

Data Availability

Table 1: Overview of data storage location.

Project	Chapter	Type of data	Location data
FRET challenge	3.2	data finder, raw data single-molecule, raw data eTCSPC	P:\FRET challenge (II)\Measurements
Calibration workflow	3.3	data finder, raw data eTCSPC, raw data sinlge molecule	P:\DNA_calibration\Measurements
Solvent relaxation	3.4	raw Data eTCSPC, measurement scripts, analysis scripts, AV settings, structural models	P:\Solvent_relaxation_jf
Holliday Junction/OFA	4.2	free dye TCSPC raw data, single-molecule raw data, simulated data	P:\Collaboration_MPI_Erlangen\data
hGBP1	4.3	se-/e-TCSPC raw data, scripts, single-molecule raw data, structural, ensembles	K:\Papers\0-In Preparation\hGBP1_jf_pl\data
Tc toxin	4.4	single-molecule raw data, TIRF raw data	P:\Tc-Toxins_Raunser\data

Contents

1 Introduction

The word biomolecule generally describes all molecules that are produced by cells and living organisms. One can divide them into four major groups which are proteins, nucleic acids, carbohydrates and lipids. As they are the fundamentals of life, understanding the working principle defined by nature of for example proteins will advance the current state of life sciences [1]. Hence, it is a long-lived goal of science to understand them in terms of their function, structure and dynamics.

Resolving the structure and dynamics of biomolecules is historically approached with using experimental methods such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, cryo-electron microscopy (cryoEM) and electron paramagnetic resonance (EPR). X-ray crystallography has resolved the atomic and molecular structure of crystalized biomolecules since long times, starting with the structure of myoglobin [2]. It is the most established method to investigate the structure of a biomolecule and it can be done also time resolved [3]. NMR spectroscopy studies molecules in solution via observation of the changes of the local magnetic field of the nucleus. After excitation of the nucleus by a pulse of radio-frequency, the relaxation out of the non-equilibrium state as a free induction decay can be observed. Doing so one can obtain information about the structural environment of the excited nucleus. NMR is used to study the structure of comparably smaller molecules (typical upper limit of around 50 kDa [4]), and more recently through further developments to also study conformational dynamics and transient intermediates [5-8]. CryoEM is applied to samples cooled to cryogenic temperatures. It was recently awarded with the Nobel prize for Jacques Dubochet, Joachim Frank and Richard Henderson in 2017 for their achievement in using this technique to resolve high-resolution structures of biomolecules. Using CryoEM, it is possible to study structural particle ensembles and through recent developments in microfluidic mixing and spraying devices also time-resolved cryoEM was enabled [9]. EPR can resolve the structure of biomolecules via attaching spin labels to the molecule of interest. Through this one can study distances between the labels of a typical range of 1-7 nm and subsequently study the structure of a molecule while it is frozen or in solution. Recently, it is also applied to study conformational dynamics [10].

Furthermore, computational methods using machine learning and prior knowledge based on resolved structures stored at the protein data bank (PDB) achieved high accuracy (AlphaFold2 [11]) in predicting a structure. The prediction of a structure is based on its sequence and the obtained information expands the amount of accessible structural information of biomolecules rapidly. In addition, using bioinformatical methods like co-evolution techniques can assist in protein structure determination [12]. Also, Molecular Dynamics (MD) simulation can yield insights into possibly dynamics of biomolecules [13].

A comparably new and rapidly developing experimental technique is single-molecule (sm) fluorescence spectroscopy [14]. It utilizes fluorescence dyes attached via a flexible linker to a biomolecule. In 2014 super-resolution fluorescence microscopy was awarded with the Nobel prize in chemistry. It has the great advantage that it can be measured in solution under near physiological conditions and track the molecule in real time. Adding more than one fluorophore arises the possibility to measure the dipolar coupling between them by Förster Resonance Energy Transfer (FRET) [15]. Using this effect, measurements can be performed within biomolecules to derive distances with Ångström accuracy [16] in a typical range of 3-12 nm. Using FRET measurements one can accurately determine structures via a combination with computational methods leading to FRET-assisted structural modeling [17]. The accuracy in the distance is achieved through detecting and combining multiple parameters using a single measurement. This

method is called Multiparameter Fluorescence Detection (MFD) and was developed in the group of Prof. Seidel [18]. The greatest advantage of fluorescence spectroscopy is the possibility to have single-molecule concentrations. This avoids any ensemble averaging and molecules can be observed individually, which offers the potential to follow the dynamics of biomolecules in real time while they perform their mechanisms. However, since molecules have to be labeled with two different dyes, labeling could be incomplete, an issue which was solved with further experimental approaches like Pulsed Interleaved Excitation (PIE) [19].

As the usage of single-molecule FRET (smFRET) studies to investigate the structure and dynamics of biomolecules is not a well-established method yet, standardized experimental procedures are missing. Most of the experiments are performed on custom-built experimental setups and a uniform way of analysis is missing. This is especially relevant for the determination of calibration factors which are critical parameters depending on the used optical devices to determine the FRET derived distances [16] and therefore conformational dynamics of biomolecules accurately.

The first main chapter of this thesis (Chapter 3) will test the reliability of the single-molecule fluorescence spectroscopy method and will further improve it in terms of a higher accuracy. A previous worldwide study tested the reliability of FRET derived distances on quasi static systems such as a double stranded DNA [16]. This test for reliability over multiple laboratories will be extended to more complex systems (proteins) and furthermore to challenging analyses like the detection of their dynamics. It turned out that setup alignment and sample calibration have the highest influence on accuracy, which led to a follow up study presented in this chapter that presents a uniform workflow for an accurate calibration of smFRET experiments. In the last section a new idea to reduce the artifact caused by a dye sticking to the surface of the protein is introduced.

The application of accurate smFRET studies can answer multiple open questions for biologically relevant samples. For a DNA-Four Way Junction (Holliday Junction, HJ) the existence of an intermediate state between two known stacked conformations is unknown, for the Human Guanylate Binding Protein 1 (hGBP1) the general working principle after the addition of a farnesyl moiety is not known and for a Tripartite Toxin Complexes (TcToxins) the structural mechanism fulfilling its biological function is not understood yet.

To answer these open questions, I will apply smFRET studies. The results will be presented in the second main chapter of the thesis (Chapter 4). The investigated biomolecules range from a size of 70 kDa to 1700 kDa (overview see Figure 1), and show kinetic and dynamic time constants on time scales differing by around 9 orders of magnitude (µs-hours). I investigated their conformational pathways as well as the change of it due to a perturbance of the environment or small part of their structure. In order to do so, a new optical device was developed that made it possible to study the fast-conformational exchange of a HJ that occurs under the addition of magnesium ions. To understand the working principle of the hGBP1 we studied it in its farnesylated and non-farnesylated state, resolving its mechanism in greater detail. As a final target, the Tc toxin was measured using different experimental approaches. Doing so we could resolve the pathway the molecule takes while it is performing a syringe like mechanism.

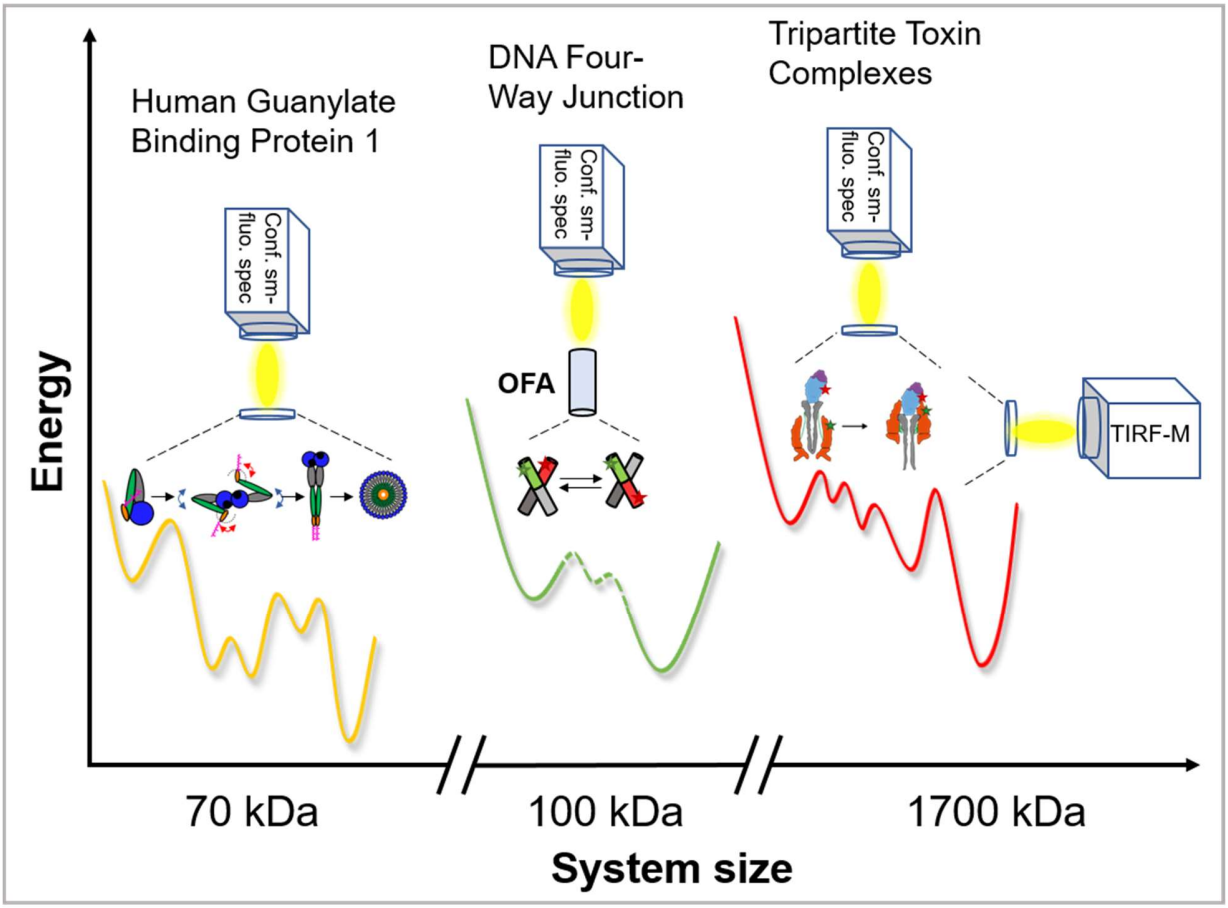


Figure 1: Target biomolecules and used method for Chapter 4. Resolved energy landscapes for a human protein (hGBP1, orange), a nucleic acid (DNA Four-way junction, green) and a toxin (Tc toxin, red). Studied molecules range from a mass of 70 kDa (hGBP1) to 1700 kDa (Tc toxin). Sketches of the methods indicate the way of measurement, where confocal single molecule fluorescence spectroscopy (conf. sm-fluo. spec.) was used as a basic principle for all investigations. In case of the DNA 4-way junction it was combined with an OptoFludic Antenna (OFA), for the TcToxin Total Internal Reflection Microscopy (TIRF-M) was also used.

2 Background – Structure and dynamics of biomolecules

2.1 Energy Landscapes

The concept of energy landscapes of biomolecules like proteins originates from the theory of the folding process of an unfolded protein into its native state. It was introduced as a solution to the "Levinthal's Paradox" [20]. The paradox can be described as follows: If one assumes that the folding process of a protein involves free sampling of all possible conformations, with each residue having at least three states, then the amount of possible conformations a 100-residue protein can possibly pass is $3^{100} = 5.1 \cdot 10^{47}$. Assuming a state stability of 1 ps for a single conformation, the time it would take a protein passing all possible conformations to find the right one would be $3^{100} \cdot 10^{-12}$ s $= 1.6 \cdot 10^{28}$ years. The universe has only existed for about $14 \cdot 10^{9}$ years, hence the paradox.

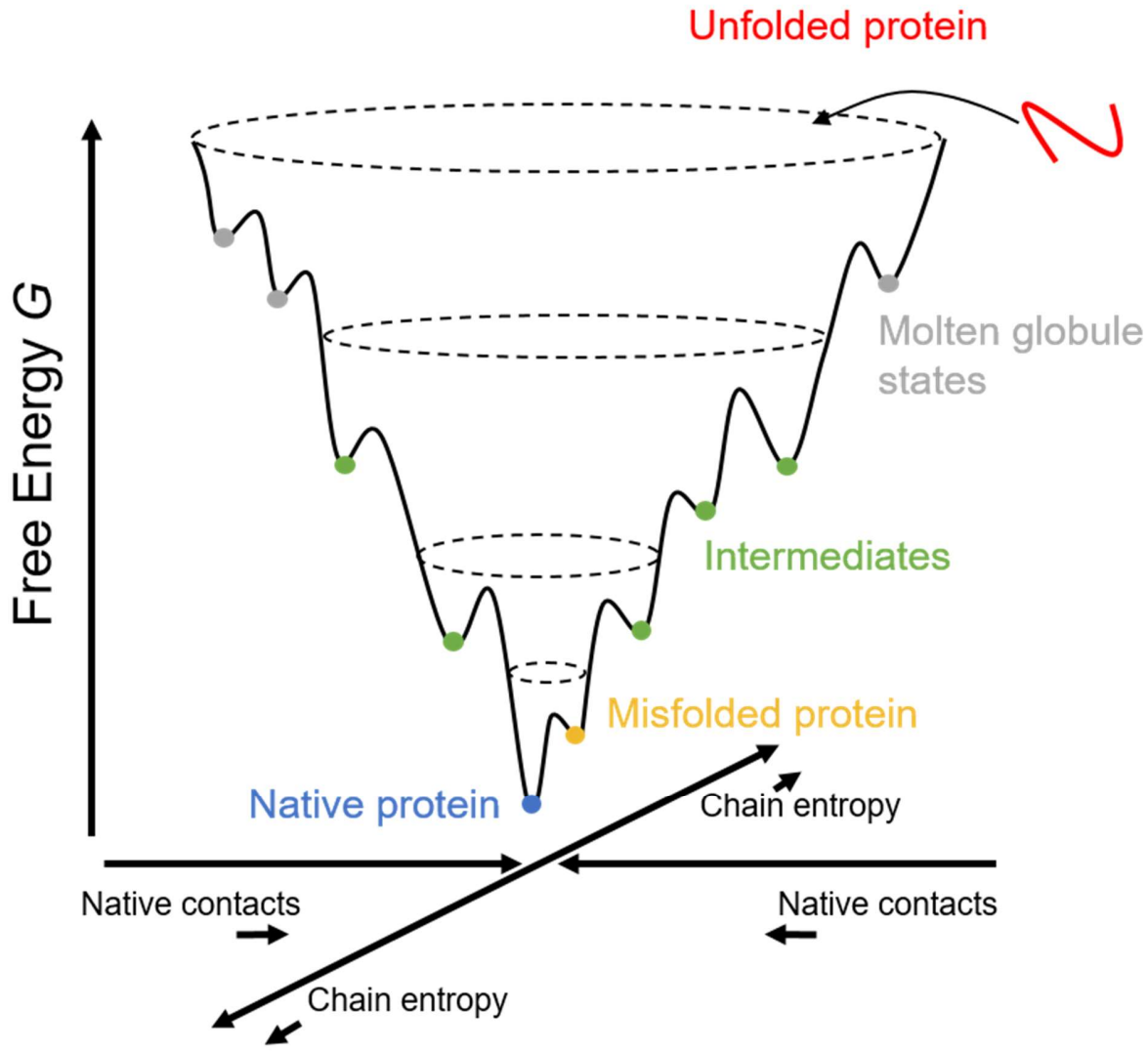


Figure 2: Energy landscape of the folding process of a protein. The molecule starts in an unfolded state (top region, red) and with a decrease of energy goes into the folded state (bottom region). Main driving force is the entropy. A protein passes molten globule states and intermediate states until it reaches its folded, native state. Figure based on [21, 22].

The answer to this paradox is that instead of the molecule passing all possible conformations, it follows a folding funnel [23], described as an energy landscape that the protein follows to find its native structure. It is energy-entropy driven and can be graphically represented (see Figure 2). Here, the molecule passes a specific state while it is decreasing both its energy and entropy. While doing so, it passes different local energy minima which are separated by energy barriers, until it finds its global minimum, which is the native state.

While protein folding is not an aim of investigation in this thesis, the same concept can also be applied to describe the dynamical behavior of folded proteins [24, 25]. Here, energy barriers in Gibbs free energy *G* define the timescales of dynamic processes (see Figure 3). Depending on the energy difference, the rate for the conformational interchange from one state to another is on a ps-timescale for small energy barriers representing for instance bond vibrations, or, for large ones, on a ms-timescale representing larger domain motions of a protein. The height of energy barriers is defined using nomenclature as introduced in previous literature [26]. This representation is only one dimensional, meaning it is a simplification of the real behavior which is highly multidimensional and depends on many interactions. The aim of this representation is to find a fundamental picture of how the biomolecule fulfills its function. The time the molecule needs to find its dynamic equilibrium is defined as the relaxation time τ_{relax}. It can also be described by the inverse of it which is a rate k. The difference of the height of for example two states of a molecule defined by local minima in the energy landscape is proportional to the equilibrium distribution of the involved states of the molecule. The height of the energy gap between them is proportional to the observed relaxation times and rates. The rates can be described by the Arrhenius equation following

$$k = Ae^{\frac{-E_a}{k_B T}} \tag{2.1-1}$$

with *A* a temperature dependent pre-exponential factor, E_a the starting energy for the chemical reaction, *T* the temperature and k_B the Boltzmann constant.

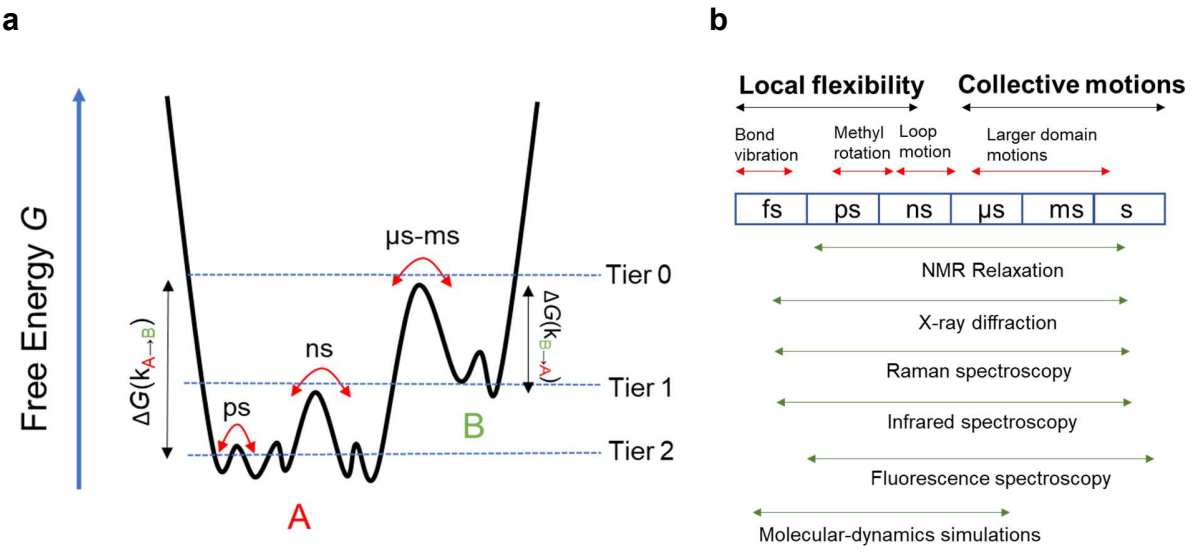


Figure 3: Equilibrium fluctuations Energy landscapes of a biomolecule defining its dynamics. a 1-dimensional energy landscape with energy barriers defining the rate constant of the dynamic equilibrium between two states (red arrow). Region of dynamics is separated in different tiers described in literature

[26]. **b** Timescale of dynamic processes in biomolecules and experimental methods that can access these timescales. Figure and part of caption based on [1].

Dynamic biomolecules are very sensitive to their surroundings, for instance the salt concentration, electrostatic interactions or the change of protonation status. Varying these parameters can have dramatic consequences on the energy landscape of the biomolecule (see Figure 4 top). Similarly, binding of a substrate or forming a dimer can change the energy landscape as well (see Figure 4 bottom). In the following thesis, energy landscapes will be used to sketch and illustrate results obtained via a change in the environment of the biomolecule.

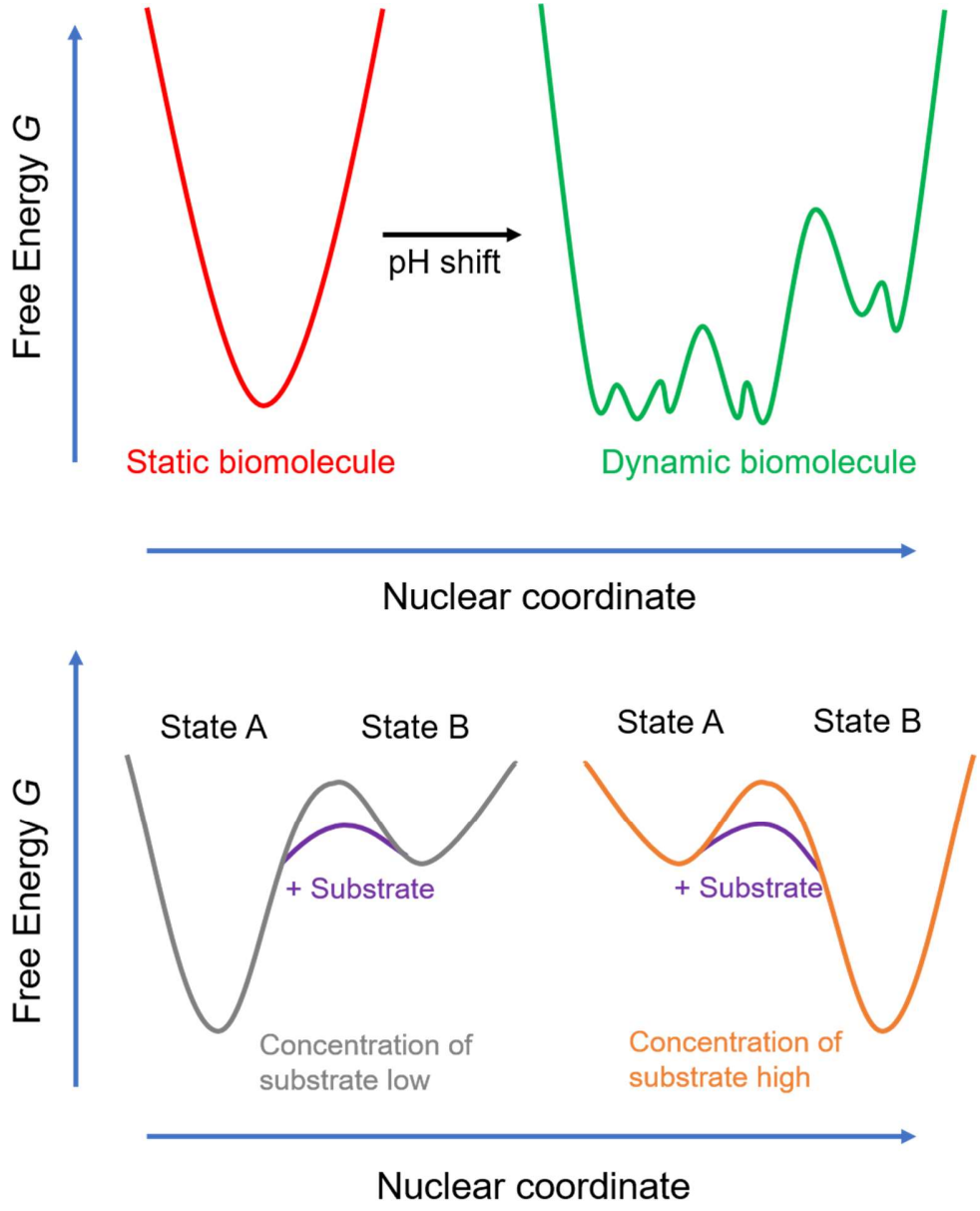


Figure 4: Modification of energy landscapes. Energy landscape can be modified by a change in the system caused for example by a pH change (top), which might possibly activate dynamics in a molecule that was static before. Another scenario would be a binding of a substrate (bottom). Here, the energy barrier to reach the other state is only passable when a substrate is bound, whereas otherwise the molecule stays

in the state A. In case of a high concentration of the substrate in the solution, the equilibrium will be shifted towards the substrate bound state (in the sketch shown as state B).

Furthermore, energy landscapes can be useful to describe a molecule undergoing a transition from one state to another (see Figure 5). It is of great interest to study the transition a molecule from a state A to another state B in real time, since it is the crucial step in the conformational equilibria. Performing these transitions, the molecule can go directly over unstable fast decaying transition states (Figure 5, left), or it can pass several more or less stable intermediates (Figure 5, right). The general way of the molecule will be defined as the transition path. Using FRET measurements to study this has also been reported in previous literature [27]. In a time dependent experiment where one is able to watch the distinct states of a molecule, these two scenarios appear in the experimental observable. While the direct transition will be seen as an instant jump of the observable from one state to the other because barrier crossing is fast, in case of going over stable intermediates, the jump is distributed in time.

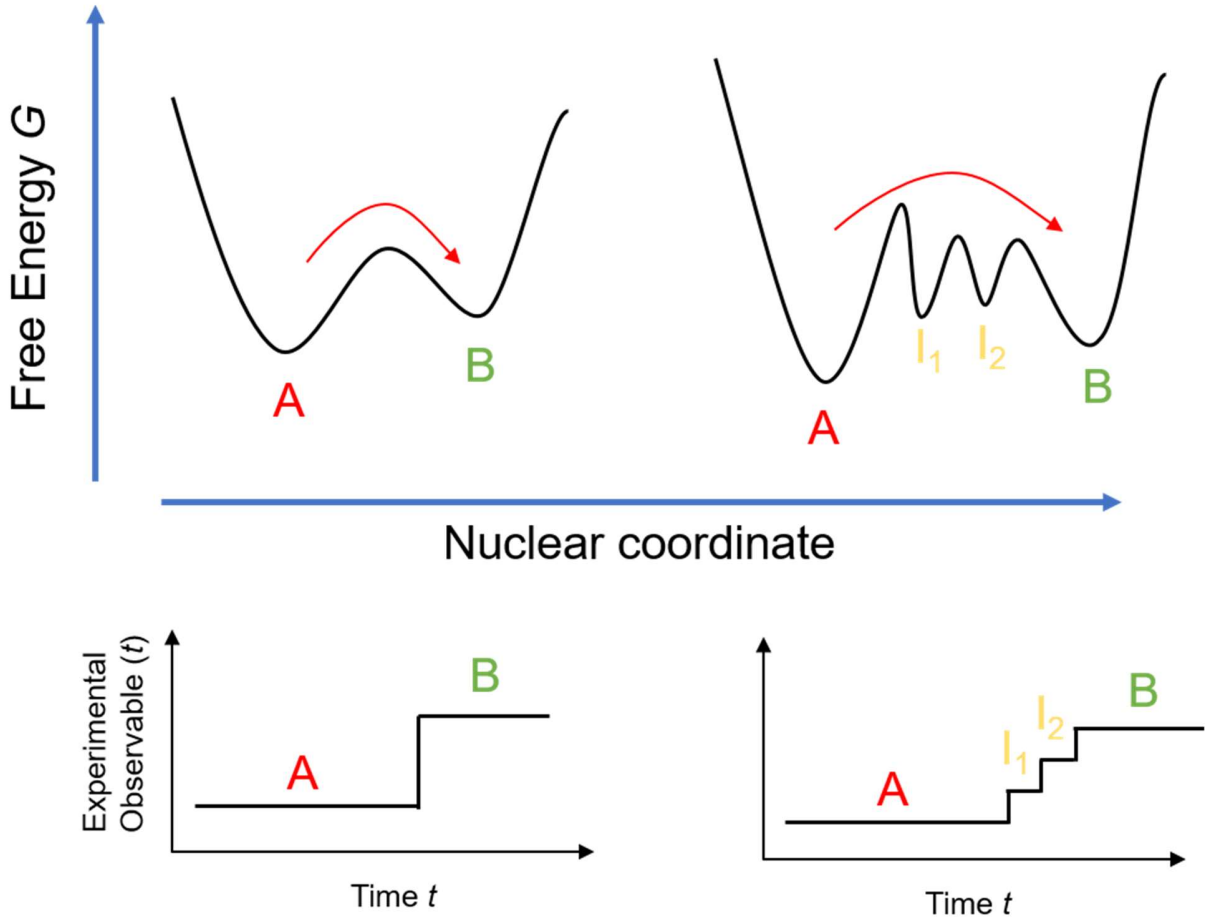


Figure 5: Schematic representation of a transition from one state to another. Top area shows the corresponding energy landscapes for a two state system going from state A to B without intermediate states (left) and with stable intermediate states (right). Bottom area shows a sketch of an experimental observable resulting from these two scenarios.

In conclusion, experimental observables like rate constants and transition path times [27] can be translated into energy landscape schemes for biomolecules. Doing so, one gets a graphical representation of the dynamical behavior of a biomolecule. In this thesis, energy landscapes are used for biomolecules of different kinds illustrating their dynamical behavior.

2.2 Biomolecular Systems

In this work, various biomolecular samples (overview see Table 2) were used to benchmark established methods, Deoxyribonucleic Acid (DNA), Maltodextrin Binding Protein (MalE), U2 Auxiliary Factor 65 (U2AF65), and to apply these methods to them, human Guanylate Binding Protein 1 (hGBP1), Tripartite toxin complexes, TcA, TcB, TcC (Tc toxin) and a DNA 4-way junction (Holliday junction, HJ). The biomolecules used in this work ranged from a mass of 21 kDa to 1.7 MDa, with various functions like immune response (hGBP1) on the one hand and infection of a cell on the other (TcToxin). A majority of them are proteins. Proteins, often phrased as "molecular machines", play many important roles in cells and tissue and are the most extensively studied subjects in life sciences. They make up to ~50% of the cell's total dry mass [28] and it is estimated that the human body contains ~100000 different proteins [29]. Another type of biomolecules, the nucleic acids are studied as well, where one (Holliday Junction) was investigated as a system itself, whereas DNA is employed as a model system with a flexibility to a distinct extent. While the expression, purification and labeling was performed in various cooperating working groups, measurement, data analysis and interpretation was done by me in the Seidel laboratory.

Table 2: Overview of biomolecular systems used in this work.

Sample	Sample short name	Type of biomolecule	Mass [kDa]	Studied in chapter	Biological relevance
Biomolecules as model systems					
U2 Auxiliary Factor 65 from human	U2AF65	protein	21.51	3.2	assembly of human spliceosomes/recognition of RNA cis elements
Maltodextrin Binding Protein from *Escherichia coli K-12*	MalE	protein	40.75	3.2	uptake and catabolism of maltodextrins
Deoxyribonucleic Acid, 36 base pairs	dsDNA	double stranded nucleic acid	22	3.3	genetic instruction
Biomolecules as aim of investigation					
DNA Four-Way Junction	Holliday Junction (HJ)	nucleic acid	98.71	4.2	genetic recombination
Human Guanylate Binding Protein 1	hGBP1	protein	68.37	3.4, 4.3	immune response
Tripartite toxin complexes, TcA, TcB, TcC from insects/human	Tc toxin	toxin/protein	1689.78	4.4	translocation of toxic enzymes into target cell membrane

2.3 Contribution of smFRET studying biomolecular dynamics

To understand the working mechanism of a biomolecule, one needs to constantly track its dynamical properties resolving every possible step the molecule might take on its pathway to a biological function. As written in a recently published article about FRET-based dynamic structural biology [30], single-molecule FRET (smFRET) measurements are an important tool in the new era of "dynamic structural biology". This is because FRET studies even at the ensemble but especially at the single-molecule level can measure structural dynamics over a minimum of 12 orders of magnitude in time [30] using ambient conditions (see Figure 6). The reason is that using smFRET studies one can measure conformational and functional heterogeneities directly. As written in a review article [1], the ultimate goal is to be able to follow the molecule in real time. To this wish, smFRET measurements can answer, as it will be shown in the scope of this thesis (see chapter 4.2 and 4.4). The main advantages of using smFRET studies to resolve conformational dynamics are the sensitivity to macro-molecular distances, measuring at single molecule level, the ability to detect rare events and the high sensitivity for labeled molecules [30].

Different kinds of experiments, techniques and analysis methods are needed to access this information leading to the resolution of various dynamic time scales. These techniques are divided into two different approaches, which use immobilized or freely diffusive molecules. Using immobilized molecules has the advantage of tracing the molecule permanently for long times, until photobleaching occurs. The techniques here are usually camera based only and access therefore slower kinetic rates. The disadvantage is usually a reduction in the signal and therefore the ability for deeper analysis. In a recent study, FRET nanoscopy was developed [31] enabling to obtain high precision FRET derived distances using immobilized samples. The other approach is the free diffusion of the molecule. Here, much higher excitation rates can be used. This leads to an increased signal, since the molecule is excited for only the diffusion time, which is typically around 1 ms. The obvious disadvantage is, that one can only see it for 1 ms in a single event, called a burst. By counting up all bursts, one can derive conformational dynamics and other useful information. A large variety of techniques has been developed to analyze smFRET experiments. Within the scope of this thesis, mainly *Efficiency-tau* plots are used [32] with recently further developed theory of FRET lines [33]. Also, Photon Distribution Analysis (PDA) [32], filtered Fluorescence Correlation Spectroscopy (fFCS) [34] and fluorescence traces, where one can follow the fluorescence of the molecule "by eye" are used.

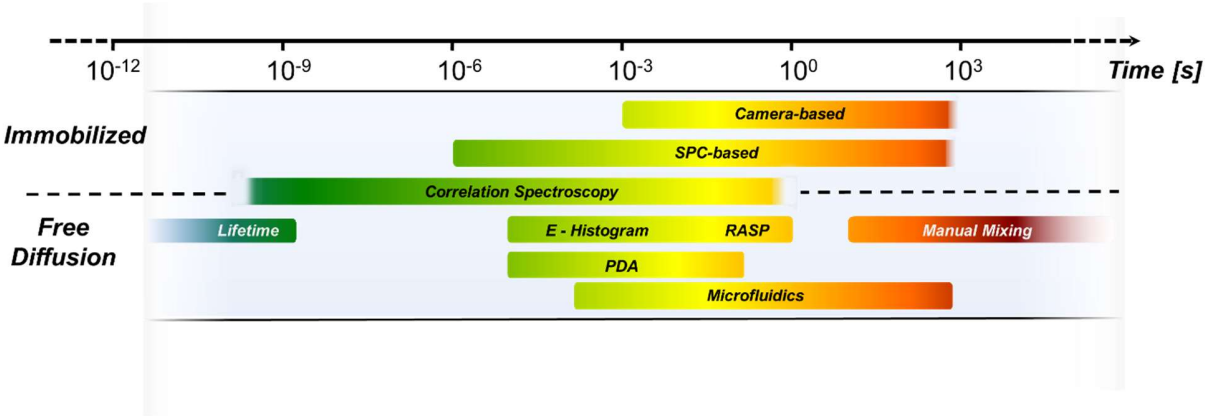


Figure 6: Dynamic range of smFRET studies. Diagram is divided into two parts indicating experimental methods using immobilized samples (top part) and freely diffusing samples (bottom part). Color code is

indicating the dynamic range the method can sense, indicating shorter times for blue/green and longer ones for red. Figure modified after [30]. Using immobilized samples one uses cameras and can combine this with Single Photon Counting (SPC). Using free diffusive molecules one can detect dynamics using the lifetime, Efficiency-Histogram (E-Histogram), Photon Distribution Analysis (PDA), Recurrence Analysis of Single Particles (RASP), microfluidics and manual mixing.

3 Methods – Benchmarking and enhancing accuracy in high precision sm-MFD

3.1 Overview

The following chapter consists of three sections. In the first section, a global FRET study is presented, in which laboratories around the world measured two different dynamic proteins. The reliability of smFRET experiments is tested with a focus on measuring FRET derived distances and dynamics in proteins. In this study, two main factors caused uncertainties in the results and difficulties in the analysis. Reason one is the uncertainty in the experimental estimation of correction factors which are needed in order to analyze the data accordingly yielding to high accuracy in the FRET derived distances. As the way of calibration can vary based on different ways of estimating correction factors or user bias, results show less agreement. Therefore, we developed an enhanced, uniform workflow to calibrate a setup and sample accordingly. This workflow will be presented in the second section. The other difficulty in the global FRET study was that dye pairs were used at positions which caused a high sticking of the dye to the surface of the protein, leading to unwished effects on the data. In order to approach this issue, I introduce an idea based on the effect of solvent relaxation in the third section of chapter 3.

3.2 How reliable and accurate is single-molecule FRET for characterization of structural dynamics and distances in proteins?

Christian Gebhardt[1,#], Ganesh Agam[2,#], Milana Popara[3,#], Rebecca Mächtel[1,], Julian Folz[3], [Benjamin Ambrose[4], Neharika Chamachi[5], Timothy D. Craggs[4], Marijn de Boer[6], Dina Grohmann[7], Taekjip Ha[8], Andreas Hartmann[5], Jelle Hendrix[9,10], Verena Hirschfeld[11], Christian G. Hübner[11], Thorsten Hugel[12,13], Carine Jackers[10], Dominik Kammerer[14], Hyun-Seo Kang[15], Achilles Kapanidis[14], Georg Krainer[5], Kevin Kramm[7], Edward Lemke[16,‡], Eitan Lerner[17], Emmanuel Margeat[18], Kristen Martens[19], Jens Michaelis[20], Jaba Mitra[21], Gustavo G. Moya Muñoz[1], Robert Quast[18], Nicole Robb[14], Michael Sattler[16], Michael Schlierf[5], Jonathan Schneider[1], Tim Schröder[2], Anna Sefer[20], Piau Siong Tan[15], Johann Thurn[12], Philip Tinnefeld[2], John van Noort[19], Shimon Weiss[22,23], Niels Zijlstra[1]], Anders Barth[2,3,†,*], Claus A. M. Seidel[3,*], Don C. Lamb[2,*], Thorben Cordes[1,*]

[1] Physical and Synthetic Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, Großhadernerstr. 2-4, 82152 Planegg-Martinsried, Germany

[2] Department of Chemistry, Ludwig-Maximilians-Universitaet Muenchen, Butenandtstr. 5-13, 81377 München, Germany

[3] Molecular Physical Chemistry, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

[4] Department of Chemistry, University of Sheffield, S3 7HF, UK

[5] B CUBE – Center for Molecular Bioengineering, TU Dresden, Arnoldstr. 18, 01307 Dresden, Germany

[6] Molecular Microscopy Research Group, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

[7] Department of Biochemistry, Genetics and Microbiology, Institute of Microbiology, Single-Molecule Biochemistry Lab, University of Regensburg, Regensburg, Germany

[8] Materials Science and Engineering, University of Illinois Urbana-Champaign, Urbana IL 61801, USA

[9] Dynamic Bioimaging Lab, Advanced Optical Microscopy Center and Biomedical Research Institute, Hasselt University, Agoralaan C (BIOMED), B-3590 Hasselt, Belgium

[10] Department of Chemistry, KU Leuven, Celestijnenlaan 200F, B-3001 Leuven, Belgium

[11] Institute of Physics, University of Lübeck, Germany

[12] Institute of Physical Chemistry, University of Freiburg, Germany

[13] BIOSS Centre for Biological Signaling Studies, University of Freiburg, Freiburg im Breisgau, Germany

[14] Department of Physics, University of Oxford, Oxford OX1 3PU, UK

[15] Biomolecular NMR Spectroscopy, Department Chemie, Technische Universität München, Lichtenbergstr. 4, DE-85747 Garching, Germany

[16] Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany

[17] *Department of Biological Chemistry, Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem, 91904, Israel*

[18] *BS, CNRS, CINSERM, Université de Montpellier, Montpellier, France*

[19] *Biological and Soft Matter Physics, Huygens–Kamerlingh Onnes Laboratory, Leiden University, Leiden, The Netherlands*

[20] *Institute for Biophysics, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany*

[21] *T.C. Jenkins Department of Biophysics and Program in Molecular Biophysics, Johns Hopkins University, 3400 N. Charles St, Baltimore MD 21218*

[22] *Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA*

[23] *California NanoSystems Institute, University of California, Los Angeles, CA 90095, USA*

[‡] *current address: Biocentre, Departments of Biology and Chemistry, Johannes Gutenberg, University Mainz and Institute for Molecular Biology gGmbH, 55128 Mainz, Germany*

[†] *current address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands.*

[#] *these authors contributed equally to this paper*

[*] *corresponding author emails:* anders.barth@hhu.de, d.lamb@lmu.de, cseidel@hhu.de, cordes@bio.lmu.de

[] authors within brackets are listed alphabetically

Manuscript stage: BioRxiv (https://doi.org/10.1101/2022.08.03.502619), in review at Nature Methods

Current version of the manuscript attached under chapter 7.2.

Author Contributions: C.G. and T.C. initiated the study. C.G., G.A., A.B., C.A.M.S., D.C.L. and T.C. designed research. A.B., C.A.M.S., D.C.L. and T.C. supervised the project. R.M. cloned and purified MalE. H.S.K. and M.S. provided U2AF65. C.G. and G.A. performed labeling of MalE and U2AF65 variants for shipment to participating labs. M.P., J.F. and C.A.M.S. designed MalE mutants 4 and 5 in silico. M.P. and J.F. performed initial measurements on MalE mutants 4 and 5. G.A. reperformed the analysis on the provided raw data for U2AF65 and MalE-1 from 8 laboratories. G.A., A.B. and M.P. performed dynamic shift estimation. C.G. performed FCS experiments on MalE variants, time-resolved anisotropy experiments and R_0-determination. M.P. performed time-resolved anisotropy analysis of single labelled MalE cysteine mutants from ensemble measurements as well of all MalE and U2AF65 dye combinations from smFRET measurements. C.G., G.A. and M.P. performed measurements of MalE mutants with additional dye combinations and M.P. and A.B. performed statistical analysis of dynamic shifts and anisotropies. G.A. and A.B. performed estimation of setup-dependent parameters and PDA of U2AF65. G.A. performed the filtered-FCS and A.B. performed TCSPC analysis of U2AF65. C.G., G.A. and M.P. performed smFRET measurements on dsDNA rulers. M.P. performed AV and ACV modelling of dye distributions for MalE and U2AF65. G.M. performed MST experiments. M.d.B. performed confocal scanning experiments for surface-immobilized MalE. All authors were involved in performing comparison experiments and analyzing smFRET data. C.G. and G.A. consolidated data collection of participating labs. C.G. and G.A. designed Fig. 1. C.G., G.A. and M.P. designed Fig. 2. A.B. designed Fig.s 3 and 4. A.B. and M.P. designed Fig.s 5 and 6. C.G., G.A., M.P., A.B.,

C.A.M.S., D.C.L. and T.C. interpreted data and wrote the manuscript in consultation with all authors.

3.2.1 Aim of the study

SmFRET experiments are an excellent tool to study the dynamics of biomolecules. In a recent blind study of double stranded DNA [16] a high reproducibility of different laboratories around the world was achieved for FRET derived distance with an interdye distance precision of ≤ 0.2 nm leading to a relative value of around 5 %. Here, the next step is taken, which is a new worldwide study of more challenging samples, which entails biomolecules like proteins. These samples are more difficult to handle due to degrading processes, having more complex local chemical environments affecting the dyes, and on top they can potentially undergo conformational dynamics. As model systems, maltose-binding protein, MalE and U2 auxiliary factor 65, U2AF65, were used (see Figure 7). These proteins were measured from various laboratories challenging the reliability of smFRET measurements. First, MalE was measured since it shows simpler conformational dynamics. Here, the focus was to test FRET derived distance changes based upon a conformational change. For the U2AF65 protein the dynamical behavior is more complex and shows an ensemble of different structures and equilibria. The following summary is therefore divided into two sections, where the first deals with the reliability of distances and the second concentrates on the reliability of analyzed dynamics.

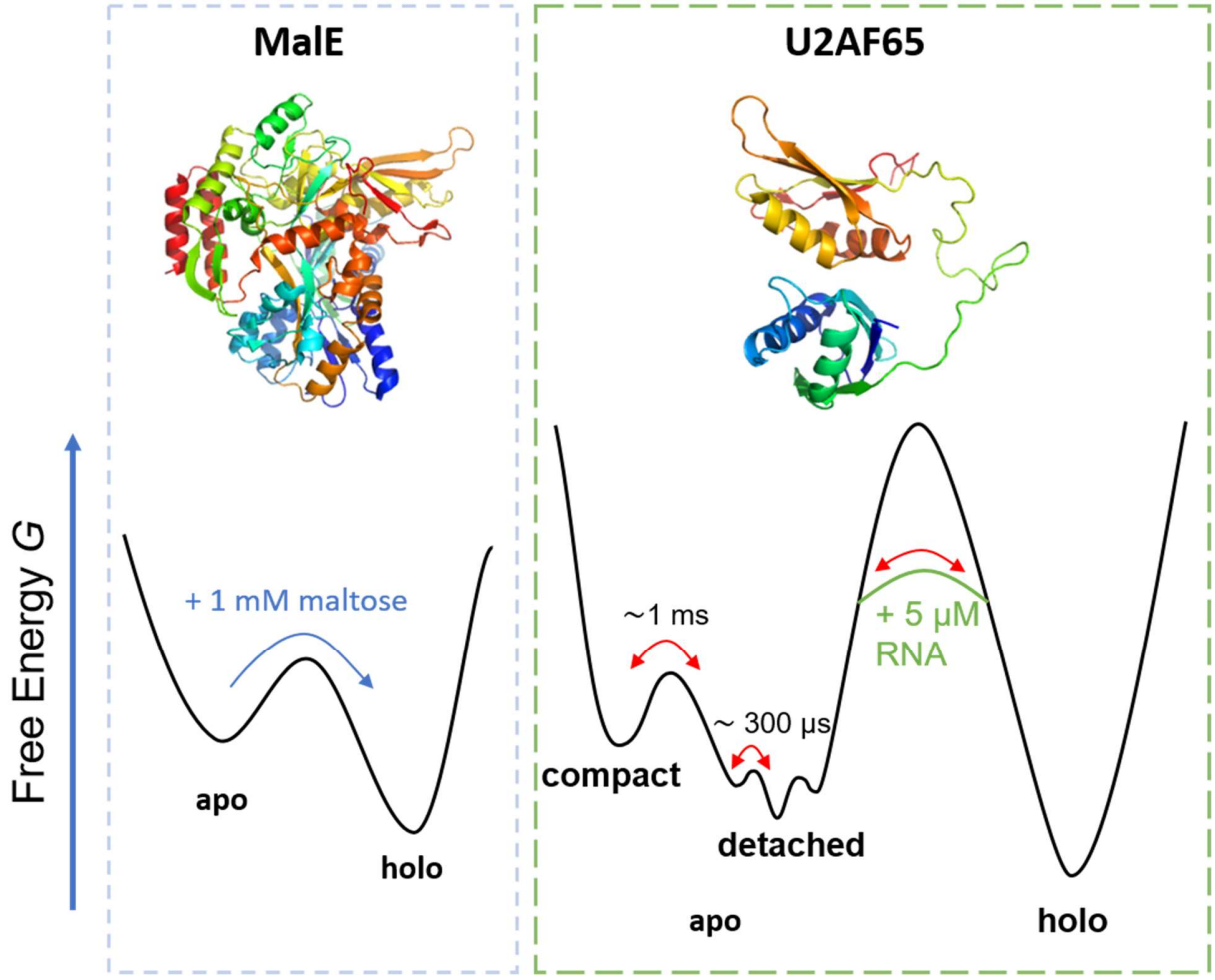


Figure 7: Energy landscapes of the two model systems for the global smFRET study. The blue box is showing the MalE protein (PDB-ID: *1omp*) and the green box the U2AF65 (PDB-ID: *2YH0*) protein. For MalE, after adding 1 mM of a substrate which is in that case maltose, the molecule is under saturated ligand concentrations and exclusively in the holo state. For the U2AF65, adding RNA leads to an equilibrium of apo and holo states.

3.2.2 Key results

Reliability of FRET derived distances of proteins

For MalE three different FRET samples with different label positions were chosen to be measured by all participating labs (N = 16). These samples are MalE-1 (K29C-S352C), MalE-2(D87C-A186C) and MalE-3 (A134C-A186C) labeled using Alexa Fluor 546 as the donor and Alexa Fluor 647 as the acceptor dye. They differ in their initial FRET efficiency as well as the effect based upon ligand binding, which is maltose. After adding maltose, the protein undergoes a structural conformation transition which leads to an increase of the FRET efficiency for MalE-1, to a decrease for MalE-2, and no change at all for MalE3. As measurements were done under saturated ligand concentrations (1 mM maltose, $K_d \approx$ 1 µM), no conformational changes were expected during the measurement time, resulting in pure apo states without maltose and holo states with maltose. Therefore, this global experiment works as a test for the reliability of FRET derived distance measurements similar to the global experiment performed previously on DNA

samples [16], with presumably no dynamics at all, i.e., a quasi-static protein. The difference being that proteins are much harder to handle since they can potentially degrade and, on top, they show a chemical environment which is not as homogenous as it is in the case of a DNA sample. Therefore, the behavior of the dyes can be potentially affected. The result of this global study can be found in Figure 8. There, the measured mean FRET efficiency over all labs shows a standard deviation between $\pm$0.02 and 0.06. This leads to a precision of 0.2 nm and an accuracy of 0.5 nm against structural models, similar to the one obtained from the global study using DNA samples. It can be seen that this spread can be reduced by a factor of ~ 3 if one focuses on the FRET efficiency changes ($\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$) (see Figure 8d). This indicates that systematic errors are present, which are assumed to be the precise estimation of correction factors. Here, the γ-factor is the most difficult to determine, while also having the largest effect on the measured efficiency/distance.

In the case of U2AF65 a previously researched variant [35, 36] is chosen and labeled at L187C-G326C with Atto532 and Atto643. In contrast to MalE this protein displays conformational dynamics within the time scale of the diffusion. Thus confocal smFRET measurements can detect them. Here, only a subset of 8 groups participated and their results can be found in Figure 9. The protein shows a broad ensemble of conformations in the apo state, having a dynamic exchange of a defined, compact state (Figure 9a, dashed box, bottom structure) with an ensemble of detached structures. Therefore, all groups found a broad population with an average FRET efficiency of $E = 0.75 \pm 0.05$ (see Figure 9b). An individual overlay of the distribution of FRET efficiencies indicates that the fractions of the states varied from different laboratories. Since protein kinetics are sensitive to the environmental conditions, it is assumed that different temperatures in the laboratories are accountable for this effect. As some outliers appeared, data was reanalyzed by a single person and an improvement of the agreement in FRET efficiencies E was observed resulting in a standard deviation of $\pm$0.02 (see Figure 9d). This indicates that besides other effects as dye artifacts, user bias can affect the results. Adding 5 µM RNA, a second narrower peak appeared, being the quasi-static holo state of the protein, because of a slow exchange of the RNA bound state, not visible in confocal smFRET measurements (see Figure 9c, d). Here, a difference to MalE is that not all molecules go into the RNA bound state, a significant fraction (15 %) of molecules stays in the apo state for the used RNA concentration.

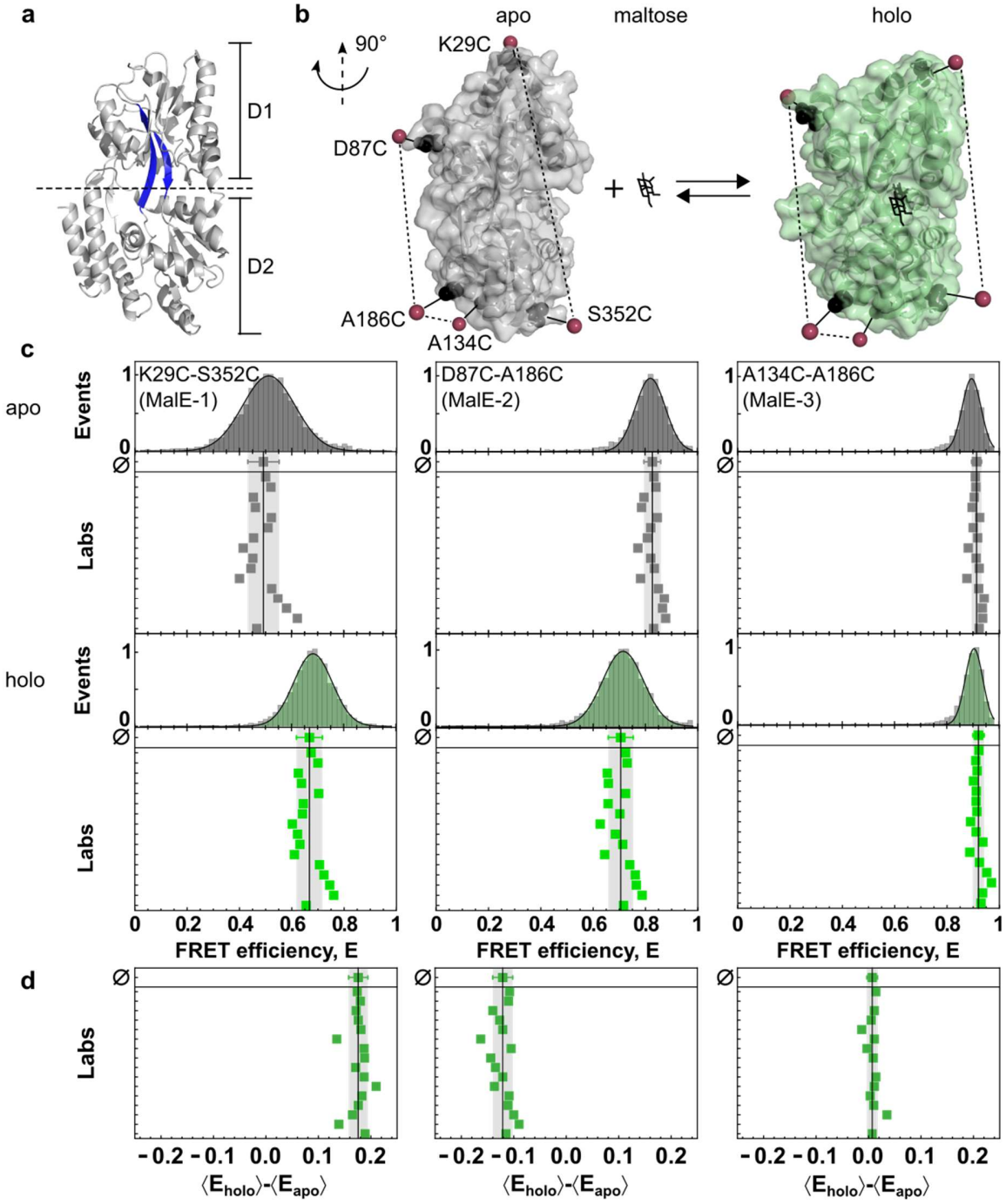


Figure 8: Global experiment of MalE using smFRET measurements. a Crystal structure of the ligand free apo state (PDB 1OMP) divided into two domains D1 and D2, which are linked by a flexible beta sheet marked in blue. **b** MalE in apo (left) and holo (right) state after ligand binding. Red spheres are indicating mean positions of the labeled dyes at different positions for all three samples, which are MalE-1 (K29C-S352C), MalE-2(D87C-A186C) and MalE-3 (A134C-A186C). **c** Exemplary FRET efficiency histograms (top) in gray for the apo and in green for the holo state of MalE. Below this the mean value for every participating laboratory is shown with a black line indicating their mean value. **d** Measured difference in the efficiency between apo and holo state for each laboratory.

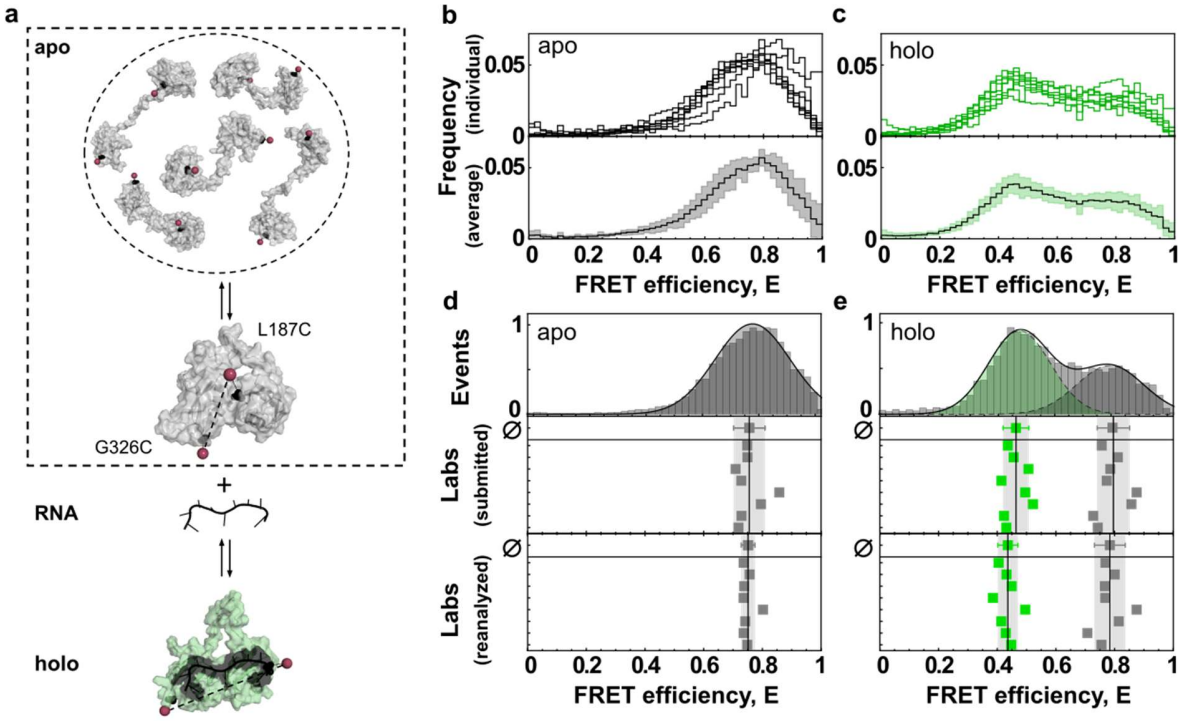


Figure 9: Global experiment of U2AF65 using smFRET measurements. a Schematic representation of the dynamic behavior of the protein. In the apo state (dashed box) the molecule is in a fast exchange between an ensemble of detached structures represented by 5 different structures (dashed sphere) and a compacted conformation (PDB: 2YHO). The holo state reached after binding an RNA molecule is shown in green (PDB: 2YH1). It is a compact but open conformation. Mean positions for attached dyes at L187C and G326C are shown by red spheres. **b, c** FRET efficiency histograms reported by the participating labs (*N* = 8) for apo (**b**) and holo (**c**) measurements. Top shows an overlay of individual distributions, bottom the average of all (black line) including the standard deviation (light area). **d** (Top) Representative FRET efficiency histogram from lab 1 for the apo measurement fitted using a Gaussian distribution with a mean value of $E = 0.75 \pm 0.05$. (Middle) Reported mean FRET efficiency values for every lab as reported and (bottom) analyzed by a single person. **e** Same representation as in (**d**) for the holo measurement. Here, two Gaussian distributions were used since not all molecules show the RNA bound state.

Reliability of dynamics detected by smFRET measurements

The detection of conformational dynamics of proteins can be achieved in diffusive smFRET experiments via various analysis methods. These are mentioned in the manuscript (see 7.1) and also in chapter 2.3 and in great detail in recently published literature [30]. In our study, most used techniques were burst-variance analysis (BVA) [37] and FRET efficiency versus fluorescence-weighted donor lifetime analysis (*Efficiency-tau* plots) [32, 33, 38]. BVA is a statistical analysis of the deviation of the FRET efficiency of a single burst and comparing it to the shot noise. In case of a dynamic system, the standard deviation will be higher than the expected shot noise limited one for a static system, leading to a dynamic shift [37]. In this summary, only the results from *Efficiency-tau* plots are discussed for simplicity. Here, a dynamic shift (*ds*) is observed from the static FRET line [33] due to the bias towards longer lifetimes since the majority of photons are emitted by them. Hence, for a mixture of species in the sample the observed FRET efficiency, which is determined via intensity and species averaged is not equal to the quantity lifetime, which is fluorescence averaged, because only a single lifetime can be determined from smFRET data. To quantify this, a dynamic shift ds is used, which is defined as the shortest distance of the

potentially dynamic FRET population to the static FRET line (see Figure 10a, blue arrow) (further details can be found in the manuscript (see 7.1). This dynamic shift could be clearly seen by every group for the apo measurement of U2AF65 and a high consistency in the data analysis was found (see Figure 10b and manuscript, see see 7.1), while MalE-2 apo seems to be static without any dynamic shift. However, for a few participating groups a dynamic shift is measured and reported (see Figure 10c). Especially for MalE-1, the measured dynamic shifts exceed the expected range in comparison to a known static system of a DNA and even to the expected dynamic shift for the apo-holo conformational change (see Figure 10c). For further investigation, the dye behavior was studied in greater detail. Additional mutants (MalE-4 and MalE5) with a higher FRET contrast for the apo-holo transition were designed and new dye pairs were used (Alexa546-AbbSTAR635P, Atto532-Atto643 and Alexa488-Alexa647). To assess the dye pair quality and to filter out dye artifacts, the combined residual anisotropy was used, defined as $r_{c,\infty} = \sqrt{r_{\infty,D} \cdot r_{\infty,A}}$. For finding a threshold value upon one could justify proper dye behavior, the effect of the residual anisotropy value was used to calculate distance uncertainties based upon the "diffusion with traps" model by S. Kalinin. The result is that a $10\,\%$ distance uncertainty results in a threshold of $r_{c,\infty} < 0.25$. It could be seen that especially for MalE-1 and the in the global experiment used dye pair, an anisotropy above a certain threshold is found (greater details can be found in the manuscript under see 7.1). An analysis over all used samples and dye pairs shows a correlation between the measured dynamic shift and the combined residual anisotropy (Pearson's *r* = 0.73). To reduce this effect, only dye pairs passing this threshold were used.

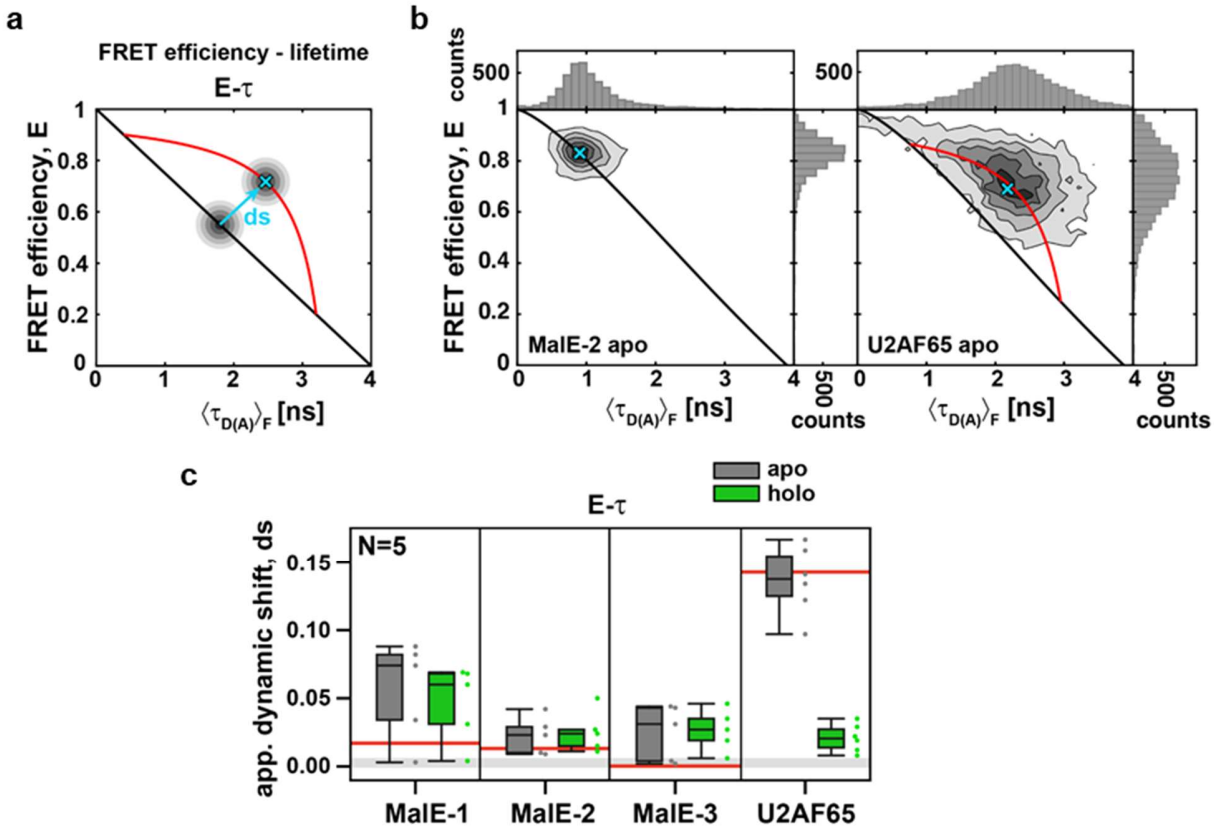


Figure 10: Detection of conformational dynamics for MalE and U2AF65 using *Efficiency-tau* plots. a In case of no dynamical averaging through conformational changes the FRET population is on the static FRET line (black line). In case of dynamics this population is shifted to the dynamic FRET line (red line).

The shift is towards longer fluorescence weighted lifetimes $\langle \boldsymbol{\tau}_{\boldsymbol{D(A)}} \rangle_{\boldsymbol{F}}$, due to the higher number of photons originating from longer lifetimes. Dynamic shift (ds) is defined as the shortest distance of the mean value of the population to the static FRET line (blue arrow). **b** Representative measurement of MalE-2 apo (left) and U2AF65 apo (right). The static FRET line is corrected for the flexibility of the linker, hence the curvature towards higher efficiencies/shorter lifetimes. For MalE-2 no clear dynamic shift is visible, whereas U2AF65 apo is clearly shifted away from the static FRET line. **c** Overview of apparent dynamic shift for the samples used in the global study under apo (grey) and holo conditions (green). Boxes indicate the median and 25%/75% quartiles, whiskers extend to the lowest and highest data point within 1.5 times the interquartile range. Grey area indicates the observed apparent dynamic shift of a DNA sample. Red line illustrates the expected dynamic shift based on a transition from the apo to the holo structure of the protein, based on their structural models.

Nevertheless, the dynamic shift is still present for some MalE-samples and a comparison based on a p-test analysis between the dynamic shift for DNA rulers and protein samples showed that it is still significant. To answer this, MD-simulations were performed by M. Popara and small, inter-residue distance fluctuations, with a standard deviation of ~3 Å are found, overshooting the typical range of thermal conformational fluctuations of ~1-2 Å [39]. Hence, it is concluded that the combination of the critical dye positions/dye pairs and these small fluctuations are responsible for the dynamic shift. Nonetheless, the question of a theoretical limit of detecting dynamics was raised and discussed. To solve this, the ratio of the dynamic shift *ds* to the experimental uncertainty σ_{SEM} was studied. It was found that to detect dynamics based upon structural changes the difference in the interdye distances need to be higher than a lower limit on the order of ≤5 Å. As a further outcome, it is demonstrated that besides the photon-counting statistics also the accuracy in the estimation of the γ-factor plays a critical role. It can be seen, that a higher uncertainty in the γ-factor directly translates into incapability of sensing dynamics (see Figure 11), since only dynamics are detectable if the ratio dynamic shift/SEM (ds/σ_{SEM}) exceeds one. We denote that it was possible to resolve the small-scale fluctuations of MalE-1, even though based upon this derived theory it should not be clearly resolvable. The reason it was still resolved is that linkers of the dyes act as a lever arm, amplifying these small-scale fluctuations.

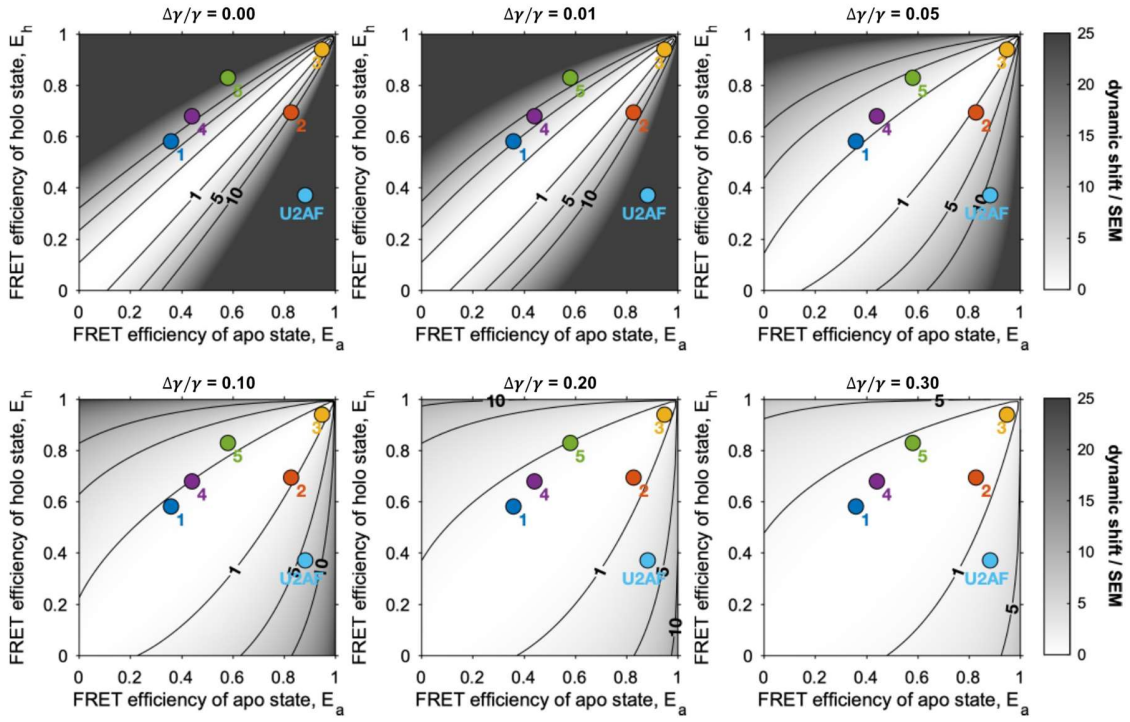


Figure 11: Theoretical limit of the detection of conformational dynamics using *Efficiency-tau* in relation the certainty in the γ-factor. Relative uncertainty of γ is given above the plot. *x*- and *y*-values correspond to the calculated efficiencies *E* based on the structural models for the apo and the holo state, respectively. For each sample (MalE-1 to MalE-5 and U2AF65), colored spheres represent these two values for each sample. Black lines indicate specific values of the ratio of the dynamic shift *ds* to the experimental uncertainty σ_{SEM}. Ratio gradient of ds/σ_{SEM} is shown as a grey scale. Dynamics are detectable if the ratio exceeds one.

3.2.3 Conclusion

It can be seen that despite of using proteins which are sensitive to outer conditions and in general more challenging systems compared to DNA, the obtained structural distance information show great agreement. This is remarkable since the involved setups and analyzing software are different and user bias is always present. Avoiding this by a global analysis of a single person increased the agreement even more. For the dynamic analysis, the used methods result in a uniform analysis for the U2AF65 apo state, which posed a challenge since it is not a simple two state system. Nevertheless, resolved relaxation times show good agreement. For the dynamic analysis in MalE a high influence of the used dye pair results in a diverse analysis if the molecule is dynamic or not. This is solved after filtering out dye artifacts and explaining the remaining dynamics with MD simulations showing small structural fluctuations that exceed the limit of thermally induced fluctuations. However, for both structural and dynamical analysis of this study, the results show that the correction factors have a large influence on the accuracy. Consequently, a follow up project is established aiming at uniforming the way for calibration by establishing a general workflow for calibration, which will be discussed in the next chapter. In addition, it is seen that trapping of a dye is critical for the analysis, which results in high anisotropy values. To explore and portray this in more detail, a small side project is introduced in chapter 3.4.

3.3 Workflow for accurate calibration of smFRET experiments

This chapter is based on a project which will lead to a manuscript written by Milana Popara and co-authors with author contributions as listed below. Here, a short summary of the current state of the manuscript can be found.

Milana Popara[1], Anders Barth[1,a], Julian Folz[1], Suren Felekyan[1], Ganesh Agam[2], Don C. Lamb[2], Claus A. M. Seidel[1]

Author list preliminary.

[1] *Institut für Physikalische Chemie, Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine-Universität, Düsseldorf, Germany*

[2] *Department of Chemistry, Ludwig-Maximilians-Universität München, Butenandtstr. 5-13, 81377 München, Germany*

[a] *Present address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, Netherlands*

Author Contributions: MP, JF, GA and SF performed measurements and data analysis. SF and AB performed software development; JF, MP and CAMS designed, ordered and managed samples. MP prepared figures and wrote initial draft of the manuscript. All authors reviewed the manuscript. AB, CAMS and DCL conceptualized and supervised the project.

3.3.1 Aim of the study

For an accurate determination of dynamics and to monitor dynamics in a biomolecular system using FRET measurements, one needs a careful estimation of correction factors. These originate from the property of optical components and dyes having a broad energy spectrum of emitted photons. In a recent, global FRET study [16] it was demonstrated that FRET derived distances can be accurately measured using a DNA as a model system. However, in the study which uses dynamic proteins (see chapter 3.2) it could be seen that for detecting dynamics, accurate correction parameters are crucial. The reason is that in case of not accurate ones, a false dynamic interpretation of the system could arise. Therefore, a more robust and reliable way of calibration is developed and tested within this project. In order to do so, a nearly static dsDNA sample with different FRET labels (see Figure 12) is used, since distances are well defined and analysis is simpler due to the absence of large-scale conformational dynamics. In the following, a workflow is introduced describing methods that were established within the scope of this project. On top, this workflow was tested for user bias via a blind study within the AG Seidel. Finally, a new, unique approach to estimate the ratio of spectral detection efficiencies crucial for the γ correction parameter is presented.

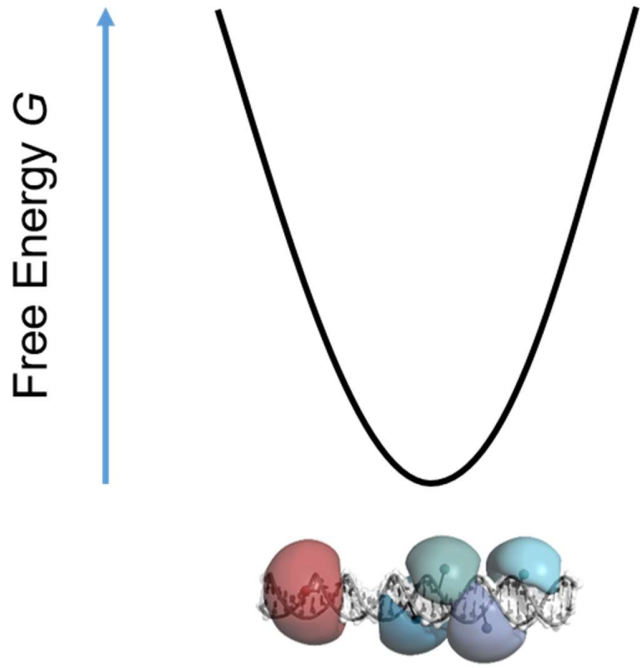


Figure 12: Energy landscape of a DNA. A DNA sample labeled with an acceptor dye (red sphere) and a donor dye at different positions (blue spheres). A DNA molecule is considered to be nearly static, given the absence of large-scale conformational dynamics.

3.3.2 Key results

General workflow and determination of correction parameters

Before one can start with the estimation of correction factors for the used dye pairs, the setup itself needs to be optimized. As a consequence, a general applicable workflow has been worked out with multiple steps (see Figure 13 top area). As most of the steps are self-explanatory, but need advanced hands-on skill, attention should be drawn to steps essential for daily calibration of the setup. These steps are for setup optimization setting accurate laser powers (2), pinhole adjustment (4) and detector adjustment (5). For setup calibration, measuring IRF (1), G-factor (2) and lifetime and anisotropy of a well-known dye (4) should be done in the beginning of every measurement. In this summary the focus will be on an analysis workflow to determine the correction parameters, although many steps are needed beforehand (see Figure 13). Diagram changed based on a figure by M. Popara et al. [unpublished work].

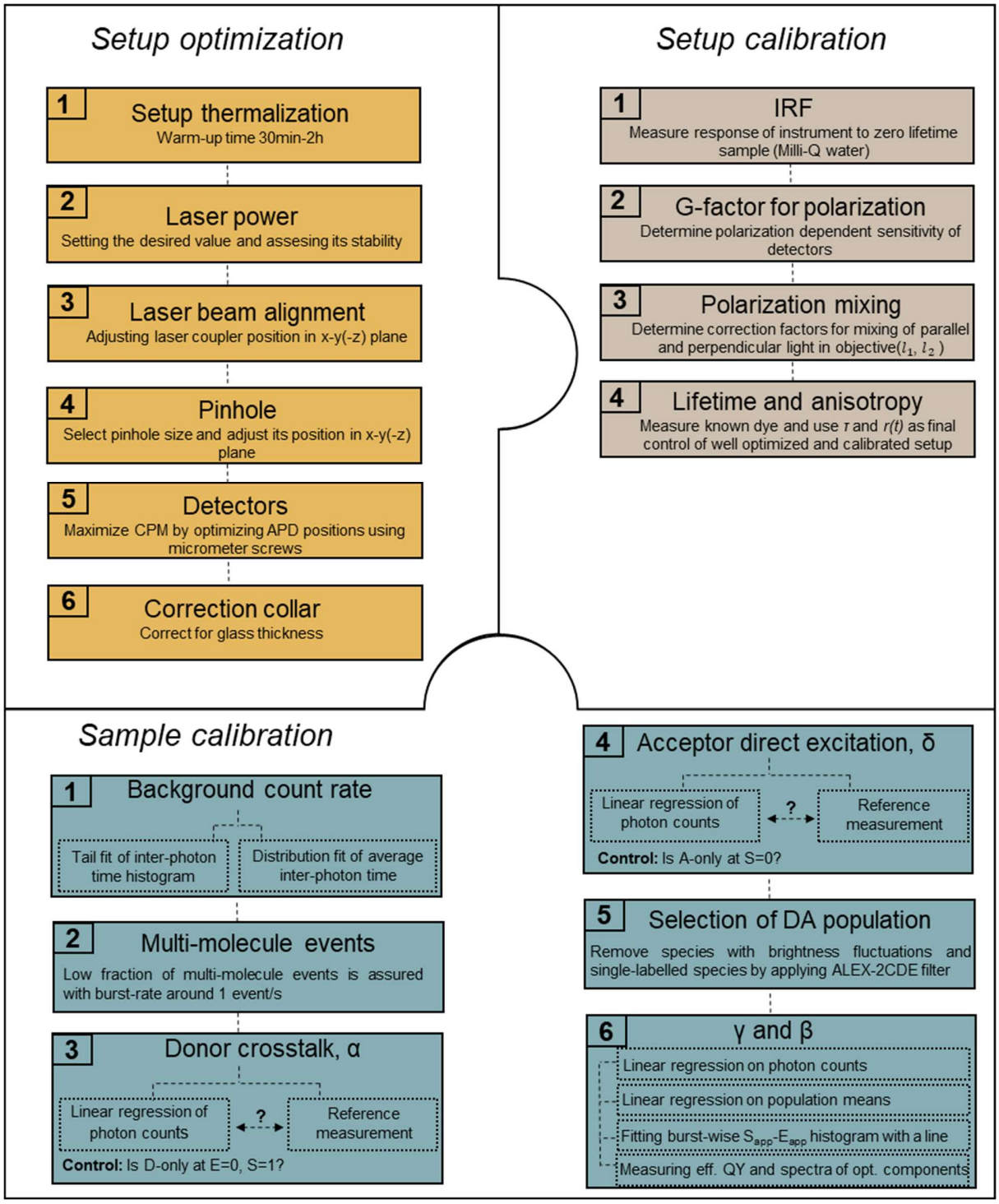


Figure 13: Diagram of workflow for measurement, sample calibration and analysis. Diagram is divided into 3 sections, one for the setup optimization (orange), for setup calibration (grey) and sample calibration (blue). Numbers and orders of panels indicates the chronological order of performing calibration. Diagram changed based on a figure by M. Popara et al. [unpublished work].

As the Figure 13 bottom area portrays, to receive a correct fluorescence signal and deduce experimental observables, one needs to correct the signal in multiple steps. The reason for that is that crucial FRET derived parameters like the FRET efficiency *E* and the stoichiometry *S* using a Pulsed Interleaved Excitation (PIE) scheme [19] are calculated as follows:

$$E = \frac{F_{\mathrm{A|D}}}{F_{A|D} + F_{D|D}} \tag{3.3-1}$$

$$S = \frac{F_{\mathrm{A|D}} + F_{D|D}}{F_{\mathrm{A|D}} + F_{D|D} + F_{A|A}} \tag{3.3-2}$$

with

$$F_{D|D} = \gamma I_{D|D} \tag{3.3-3}$$

$$F_{A|D} = I_{A|D} - \alpha I_{D|D} - \delta F_{A|A} \tag{3.3-4}$$

$$F_{A|A} = \frac{1}{\beta} I_{A|A}. \tag{3.3-5}$$

Here, *F* denotes fluorescence signal corrected via detection ratio *γ*, excitation flux *β*, spectral donor crosstalk *α* and direct acceptor excitation *δ*, and *I* denotes background corrected intensities. Indices indicate the excitation (ex) and detection (em) scheme. Therefore, accurate estimation of these correction parameters is crucial for intensity-based FRET experiments. In the following, the calculation of these parameters is chronologically introduced as shown in Figure 13 bottom, suggesting a standard workflow for sample calibration. A simplified summary of the ongoing project is presented, focusing on the workflow itself. As a side note: All introduced methods have been implemented in the software package of AG Seidel by S. Felekyan. The software package is available upon request under https://www.mpc.hhu.de/software/mfd-fcs-and-mfis.

(1) Background: As a first step, background signal needs to be estimated. Background signal obeys to Poissonian statistics and can be estimated from an analysis of the inter-photon time Δt. While background signal can therefore be fitted using a single exponential function, the fluorescence signal occurring at very short Δt cannot. Therefore, in this approach, a threshold is defined where the mono exponential decay of the background starts based on its distribution which follows:

$$P(\Delta t) = N_0 \exp\left(-S^{(BG)}\Delta t\right) \tag{3.3-6}$$

Background signal is obtained as the inverse of the exponential time constant.

(2) Multi-molecule events: In smFRET experiments multi-molecule events must be highly avoided since two bursts originating from molecules in focus at the same time cannot be separated. The probability to find m molecules in the confocal volume follows [40]:

$$P_m(N_{\mathrm{av}}) = \frac{N_{\mathrm{av}}^m}{m!}\exp(-N_{\mathrm{av}}) \tag{3.3-7}$$

with N_{av} being the average number of molecules in focus. As an example, for a typical translational diffusion time of $1\ \mathrm{ms}$ and one event per second, N_{av} = 0.001. This yields to multi-molecule events of 0.05 % ensuring that this effect is negligible.

(3) Donor crosstalk: α describes the mistaken red photons emitted by the donor dye due to its property of a broad spectrum. The ratio can be estimated from the donor-only labeled population from single-molecule measurements which is then analyzed through linear regression of the photon counts

$$N_{R|G} - S_{R|G}^{(BG)}T = \alpha\left(N_{G|G} - S_{G|G}^{(BG)}T\right). \tag{3.3-8}$$

In this notation, $N_{(X|Y)}$ denotes the photon counts in detection channel X after Y excitation, where G stands for green as the donor photons and R for red as the acceptor photons. S_{XY}^{BG} is defined as the background corrected count rate. As a control, one can check α using an additional measurement with only donor labeled molecules or donor free dyes in nM concentration using the following ratio:

$$\alpha = \frac{S_{R|G} - S_{R|G}^{(BG)}}{S_{G|G} - S_{G|G}^{(BG)}}. \tag{3.3-9}$$

(4) Acceptor direct excitation: In a similar way the direct acceptor excitation δ describes the red photons emitted by the acceptor but falsely excited by the donor excitation laser. It can be fitted using linear regression of the photon counts:

$$N_{R|G} - S_{R|G}^{(BG)}T = \delta(N_{R|R} - S_{R|R}^{(BG)}T) \tag{3.3-10}$$

and cross validated using an acceptor only labeled sample via following relation:

$$\delta = \frac{S_{R|G} - S_{R|G}^{(BG)}}{S_{R|R} - S_{R|R}^{(BG)}}. \tag{3.3-11}$$

(5) Selection of DA population: The selection of the double labeled molecules is a crucial step since one wants to discard any events from single labeled molecule in the analysis of, for instance, FRET derived distances. A robust method to do so is using the ALEX-2CDE filter [41] in combination with a stoichiometry cut, which filters out all single labeled molecules. The ALEX-2CDE filter calculates the acceptor/donor brightness ratio detected after donor excitation, $BR_{D_{ex}}$, and donor/acceptor brightness ratio after acceptor excitation, $BR_{A_{ex}}$, around each of the detected photons. It is defined as

$$ALEX-2CDE = 100 - 50 \times \left[\frac{1}{N_{CHA}}\sum_{i=1}^{N_{CHD}}\frac{KDE_{CHD,i}^{CHA}}{KDE_{CHD,i}^{CHD}} + \frac{1}{N_{CHD}}\sum_{j=1}^{N_{CHA}}\frac{KDE_{CHA,j}^{CHD}}{KDE_{CHA,j}^{CHA}}\right], \tag{3.3-12}$$

and converges to zero for a fixed brightness, whereas it increases its values for fluctuating brightness. Thus, cutting the bursts to a specific value will filter out single labeled molecule as well as bleaching effects.

(6) *γ* and *β*: Correction of these factors can be approached using different procedures:

(a) Non-linear regression of S_{app} versus E_{app}

This procedure was described in a recent global study using smFRET [16] and is based on following relation:

$${}^{iii}S_{app} = (1 + \gamma\beta + (1-\gamma)\beta\,{}^{iii}E_{app})^{-1} \tag{3.3-13}$$

where the index app refers to the uncorrected values. It is an established method but lacks a full description of the statistics of the *Efficiency-Stoichiometry* distribution. It also requires multiple FRET species and the same quantum efficiencies of each species, which is usually not the case.

(b) Linear regression on photon counts

Another approach uses photon counts instead of *E* and *S* values since they have a well-defined statistic. It follows [42]:

$$N_{R|R}^{(1)} = \gamma\beta N_{G|G}^{(1)} + \beta N_{R|G}^{(2)} \tag{3.3-14}$$

with (1) indicating a background corrected photon count and (2) a fully corrected photon count. Performing a two-dimensional robust linear regression to this equation leads to the correction factors. This has the advantage that in principle it can be applied to a single species. A drawback is that the stoichiometry needs to be 0.5 ensuring only double labeled molecules, which has to be known in advance.

(c) Linear regression on population means

As an alternative, one could fit directly the means of the donor acceptor sub-populations using the relation as described in [43]:

$$1/\langle {}^{iii}S_{app}\rangle = a + b\langle {}^{iii}E_{app}\rangle \tag{3.3-15}$$

and obtain correction factors from the intercept and slope of the linear fit:

$$\beta = a + b - 1 \tag{3.3-16}$$

$$\gamma = (a - 1)/(a + b - 1). \tag{3.3-17}$$

This method has the advantage of insensitivity to shoulders in the distribution, whereas it requires defined populations which are not too broad. Also, it is based on multiple species.

(d) Measuring and calculating γ "a priori"

Another approach is based on using additional measurements including all optical spectra of the used optical components of the setup and dyes. In principle, γ is defined as the following ratio:

$$\gamma = \frac{\Phi_{F,A}^{eff}}{\Phi_{F,D}^{eff}} \cdot \frac{g_{R|A}}{g_{G|D}} \tag{3.3-18}$$

where Φ^{eff} denotes the effective quantum yield of the dye on the sample and $\frac{g_{R|A}}{g_{G|D}}$ the ratios of the detection efficiency of the donor and acceptor dye. Here, $g_{R|A}$ is defined as the detection efficiency of a "green" (G) or "red" (R) detector for a donor (D) or acceptor (A) dye, as used in previous literature [44]. The effective quantum yields need to be measured independently or estimated using sub ensemble TCSPC. The critical point is the detection efficiency ratio. Generally, it is defined by the overlap integral of all optical components' transmission spectra to the dyes' emission spectra. Measuring and calculating this parameter will be tested using a series of free dyes in the next but one section "Comparing measured and predicted detection efficiency ratios using free dyes". This method is the most elaborated one, since it has the unique approach of knowing the parameter γ before FRET efficiency values are measured using independent measurements.

Testing workflow in a blind study

The presented workflow was tested for the influence of user bias using DNA samples labeled at different positions similar to the ones used in a recent global FRET study [16] and labeled with Alexa488 as donor- and Atto647N as acceptor-dye. For this, a data set of measurements using three different labeled samples with distances representing a low, medium and high FRET efficiency were analyzed from seven members of the AG Seidel. As a result, the distribution of obtained correction parameters using the presented workflow is shown in Figure 14. It can be observed, that the workflow leads to precise correction factors with small spread of distributions. For the most critical parameter, a standard deviation of 3 % for method (a) and (b), and 11 % for method (c) was received. Therefore, method (a) and (b) can be highly recommended for usage since they show little user bias.

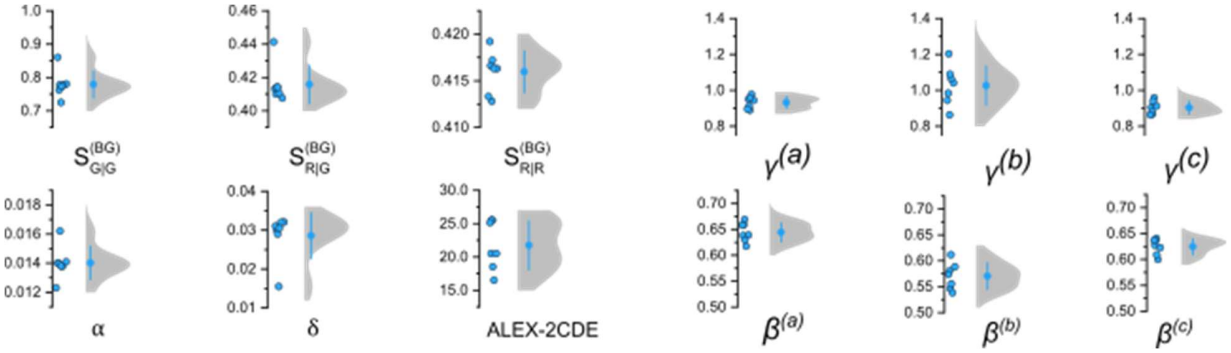


Figure 14: Blind study of the presented workflow. Values of correction parameters and distribution reported from lab members ($N = 7$) of AG Seidel. For γ-value, 3 different methods were tested (a) Fitting the burst-wise S_{app}-E_{app} histogram with a line (b) Linear regression on photon counts (c) Linear regression on population means following the suggested workflow. Figure and part of the caption taken from M. Popara et al. [unpublished work].

Comparing measured and predicted detection efficiency ratios using free dyes

As mentioned in the previous section, it is of great interest to know the detection ratio $\frac{g_{R|A}}{g_{G|D}}$ for a given setup and used dyes. As so, we tested this approach on our setup predicting this ratio for a set of dyes emitting in both green and red channels with a suitable, varying ratio and comparing it to measured ratios. For prediction all transmission spectra of each optical component were measured (see Figure 15a). For the detectors, detection quantum yields were taken from the manufacturer. Overlaying this yield to the spectral dependent instrumental detection efficiency $g(\lambda)$, which has two non-zero intervals, one in the green wavelength area, defined as $g^G(\lambda)$, and one in the red wavelength area, defined as $g^R(\lambda)$ (see Figure 15b). With the measured emission spectra of each dye j, $p_j(\lambda)$ one gets:

$$f_j^{(\Lambda)} = f_{0,j} \int_{\Lambda} g(\lambda) p_j(\lambda) d\lambda \tag{3.3-19}$$

with Λ as the sum of both λ – window for green and red detection window and $f_{0,j}$ total radiative intensity of j-th dye. Since $f_{0,j}$ is unknown and difficult to measure one uses the ratios ξ_j of the measured green (G) and red (R) λ – window, so that it is cancelled out:

$$\xi_j = \frac{f_j^{(G)}}{f_j^{(R)}}. \tag{3.3-20}$$

This quantity can be calculated as described and also measured to compare it (see Figure 15c). Furthermore, $\xi_{j,predicted}$ can be optimized to obtain better matching values for the ratio experiment/calculated. Justification to use an optimization is that the detector quantum yields are not very precise and that experimentally obtained values depend on a daily calibration of the setup. Optimization of the solution is further regularized by requiring its smoothness, since sharp changes in detection profile are not physical. Smoothness regularized optimization is implemented by Oleg Opanasyuk as in-house python script. In short, the measured set of dyes is represented as a matrix and multiplied by an instrument shape function $g(\lambda)$ as a vector with elements for G and R, and finally a function is calculated to minimalize the discrepancy between these quantities, ensuring an optimal solution for all measured dyes. Measured dyes and the ratio predicted/optimized ξ_j can be found in Table 3. As a result, optimized prediction and measured values match in very good agreement with a standard deviation of only 5 %. This is remarkable

since for this method only free dyes are used/measured which do not rely on any deeper analysis and do not depend on the quality and quantity of the sample of interest. Also, the measurement of a set of dyes is fast and does not necessarily need to be done every time before a measurement is performed, but rather on for example monthly basis using a subset of dyes to monitor the condition of the setup.

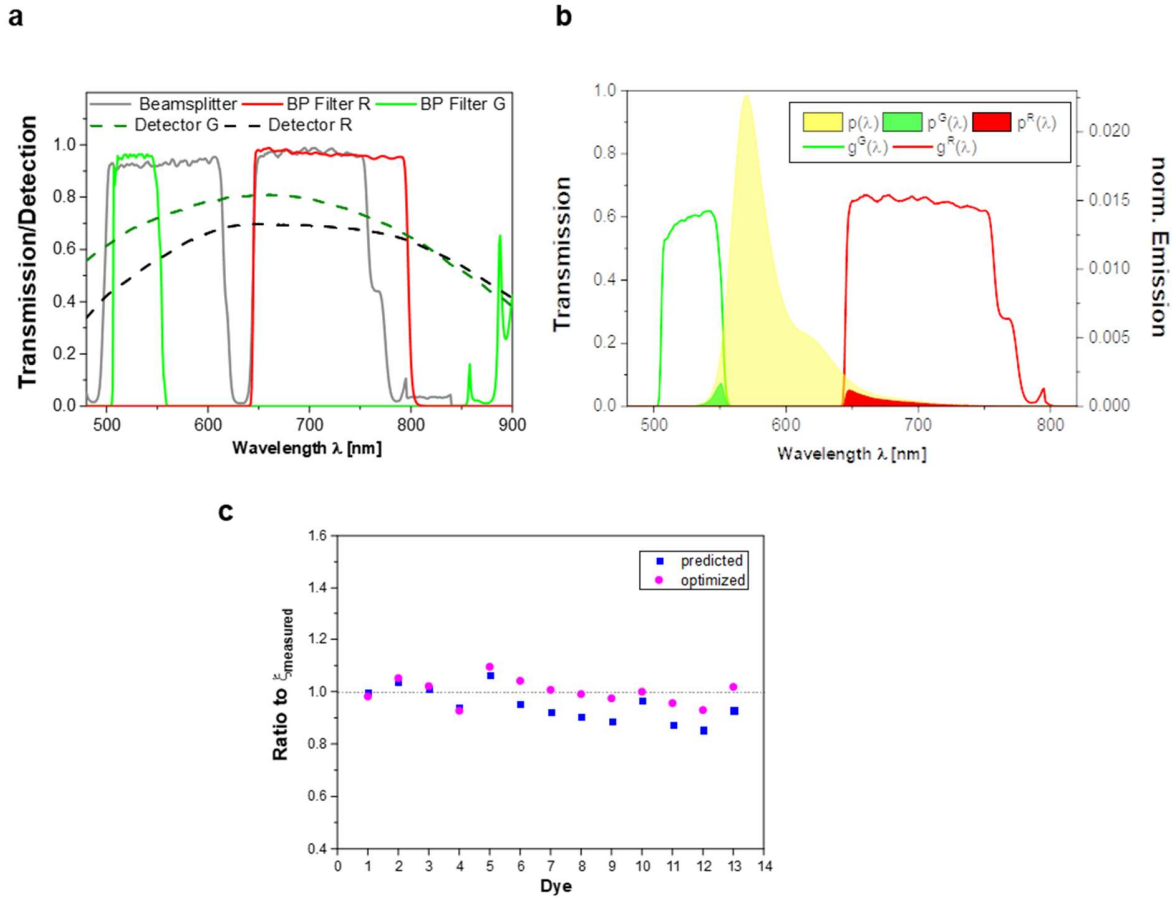


Figure 15: Measuring and prediction of green and red detection efficiency ratio. a Transmission spectra of the involved optical components which are the main dichroic beamsplitter (grey), the bandpass filter for red detection (red), bandpass filter for green detection (green) and the detector quantum yields for green and red (dashed lines). Reflection spectra for color splitter and objective transmission is left out for simplicity. **b** Overlay of the instrument shape function $g(\lambda)$ for green λ-window ($g^G(\lambda)$, green line) and red λ-window ($g^R(\lambda)$, red line) to the emission spectrum of a dye $p(\lambda)$ (yellow area). Overlap integral of p and g is shown as green transparent area for green λ-window, $p^G(\lambda)$, and in red for red λ-window, $p^R(\lambda)$. **c** Ratio of $\xi_{j,predicted}$ and $\xi_{j,optimized}$ to $\xi_{j,measured}$. Due to optimization the ratio for a dye j is closer ideal value of 1. Figure changed based on a figure by M. Popara et al. [unpublished work].

Table 3: Set of dyes used for estimation of detection efficiency ratios. Shown is the dye, manufacturer, stock solution solvent and the measured quantities which is the measured ratio $\xi_{j,measured}$ for the *j*-th dye, predicted and optimized ratio $\xi_{j,\ pre./opt.}$ and the ratio between them. Table taken from M. Popara et al. [unpublished work].

Dye	Manu-facturer	Solvent	$\xi_{j,measured}$	$\xi_{j,predicted}$	$\xi_{j,optimized}$	$\frac{\xi_{j,predicted}}{\xi_{j,measured}}$	$\frac{\xi_{j,optimized}}{\xi_{j,measured}}$
Alexa Fluor 488 free acid, 5 – isomer/ experimental sample	Molecular Probes	water	93.2809	93.1178	91.6162	0.9983	0.9822
Atto 488 – COOH	ATTO-TEC	water	69.7432	72.2228	73.3419	1.0356	1.0516
Rhodamine 110 – chloride	Sigma Aldrich	ethanol	60.5077	61.1656	61.7754	1.0109	1.0210
Oregon Green 514 carboxylic acid, mixed isomers	Molecular Probes	N, N – DMF	48.8029	45.7916	45.2581	0.9383	0.9274
Rhodamine 123	Sigma Aldrich	ethanol	40.8533	43.4413	44.7400	1.0633	1.0951
Rhodamine 19 – perchlorate	Fluka	ethanol	16.2032	15.4369	16.8769	0.9527	1.0416
Atto 532 – COOH	ATTO-TEC	water	9.4590	8.7309	9.5257	0.9230	1.0070
Rhodamine 6G	Lambda Physik	water	8.5998	7.7937	8.5223	0.9063	0.9910
Atto 542 – COOH	ATTO-TEC	water	2.8493	2.5236	2.7775	0.8857	0.9748
Rhodamine B	Fluka	water	0.5932	0.5742	0.5929	0.9680	0.9996
Alexa Fluor 546 free acid	Molecular Probes	water	0.4879	0.4268	0.4667	0.8747	0.9565
5 – carboxy tetramethylrhodamine (5-TAMRA)	Biosearch	methanol	0.3532	0.3015	0.3283	0.8537	0.9297
Atto 550 – COOH	ATTO-TEC	DMSO	0.2755	0.2560	0.2806	0.9291	1.0182

3.3.3 Conclusion

In this project, a robust workflow for determining accurate correction parameters was developed and successfully tested by a blind study within the AG Seidel. As a result, we suggest specific methods to determine background intensities, donor crosstalk, direct excitation, different excitation flux and detection ratios, which rely on different measurement parameters. Applying them is comparably simple and yields minimum user biased results. Additionally, we present a new, unique approach to measure detection efficiency ratios using a set of free dye pairs. This approach has the advantage, that it can be done before one measures the sample of interest, rendering it to an “a priori” approach. Doing so it is ensured that measuring and calibration are disentangled, which increases reliability of the data and makes certain that the calibration works in case the sample of interest does not. This workflow leads to correction factors with high accuracy and thus improves the observed difficulty in detecting biomolecular dynamics that were raised in chapter 3.2.

3.4 Using solvent relaxation to improve the accuracy of smFRET measurements

Data availability under Table 1. *Following results are based on ensemble TCSPC measurements using a FluoTime300 (PicoQuant, Berlin). Storage location of the data and scripts under* Table 1.

Data for emission and absorption spectra was taken from Alexander Larbig.

3.4.1 Aim of the study

In this short study the principle of solvent relaxation of a dye attached to a model system was measured and furthermore used to study FRET derived distances in dependence of the detection wavelength. The idea is based on the principle of solvent relaxation (see Figure 16): After excitation of the dye the orientation of its dipole moment is perturbed resulting in a change of it. Since the solvent of the dye is polar, a reorientation of the dipole moment of the solvent molecules occurs, which is on a timescale much faster than the lifetime of the excited dye and typically around 10-100 ps [45]. After that, the total energy of the system of dye and solvent loses energy due to a stabilized S_1 state compared to the case before relaxation, hence a red shift of the emitted light can be observed. In the following study, this effect was used based on the assumption that a dye which is far away from the labeled protein experiences a higher red shift compared to dyes that are trapped on the protein, which is due to the absence of solvent. To do so, energy loss of the solvent relaxation was measured to test the presence of solvent relaxation and afterwards the assumption will be tested using additional experimental observables. Finally, the effects of solvent relaxation were tested on FRET derived interdye distances. Using this feature, problems with trapped dyes which occurred in the global FRET study introduced in 3.2 can potentially be minimized. As a model system a labeled protein was used, hGBP1, which will be discussed in detail in chapter 4.3.

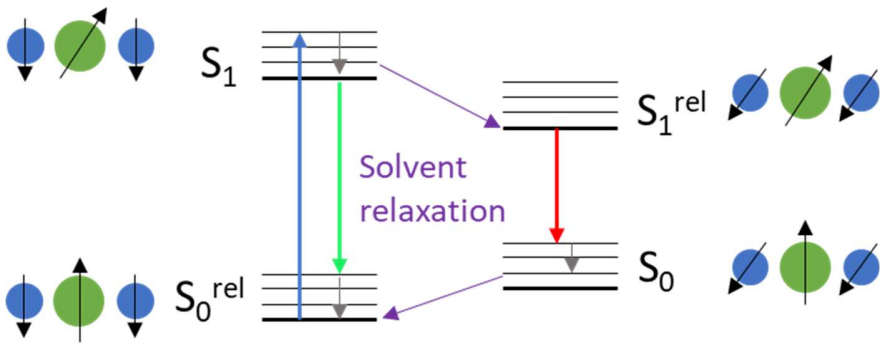


Figure 16: Jablonski diagram of an excited dye. Blue spheres represent molecules of the surrounding environment, green sphere represents the dye. Black arrows in spheres indicate the orientation of the dipole moment and its relative amplitude. Blue arrow represents the excitation, red and green arrow the emission and purple the solvent relaxation. For simplicity, only the ground state S_0 and first excited state S_1 are shown with S^{rel} indicating the relaxed state

3.4.2 Results

As a first step, the effect of solvatochromism was tested. To do so, a test system of an only donor labeled hGBP1 was used, labeled at amino acid 344 with Alexa488. Based on the emission spectrum of Alexa488, 16 different detection wavelengths λ were selected, ranging from 500 nm to 575 nm in 5 nm steps. To set a maximum specific detection wavelength, minimum slit width was chosen yielding to a $\Delta\lambda$ = 5 nm. For each detection wavelength, time resolved fluorescence decays were recorded under different detection angles, one parallel $F_{||}$ to excitation, one perpendicular $F_{\perp}$ to excitation and one under magic angle conditions eliminating any polarization effects. To measure the expected red shift, emission spectra F_{tot} were assembled out of $F_{||}$ and $F_{\perp}$ based on

$$F_{tot}(t,\lambda) = F_{||}(t,\lambda) + 2G(\lambda)\,F_{\perp}(t,\lambda) \tag{3.4-1}$$

and the anisotropy was calculated using

$$r(\lambda) = \frac{F_{||}(\lambda) - G(\lambda)F_{\perp}(t,\lambda)}{F_{||}(\lambda) + 2G(\lambda)\,F_{\perp}(t,\lambda)} \tag{3.4-2}$$

with $G(\lambda)$ as the ratio of $F_{||}/F_{\perp}$, measured using free dye not attached to a protein. In Figure 17a and b the solvatochromism of Alexa488 is tested. As it can be seen, the emission and absorption spectra are shifted depending on the polarity of the solvent. As an example, in Figure 17a the spectra in water and DMSO are shown. Here, the spectra in DMSO show a higher red shift compared to the more polar solvent water. The solvent shift, which is the shift of the emission spectrum to the absorption spectrum is plotted in Figure 17b, where one can see that for all solvents a shift between the fluorescence and absorption can be observed. Since the slope of the linear fit is unequal to 1, one expects a dependency on the polarity of the solvent. This proves a measurable effect of solvatochromism of Alexa488. The red shift of the emission can be also seen when the dye is attached to a protein (see Figure 17c). While measuring the emission spectrum a red shift occurred after a specific relaxation time. To study this, the solvent shift was measured using different relaxation times (see Figure 17d) and estimated to $\tau_{solvent\ relaxation}$ = 133 ps. Here, the energy loss in wavenumbers was calculated using the average emission wavelength. A second component occurred in the biexponential relaxation process, with a time constant of τ = 8 ns. While the short component is in approximate range of the expected time of solvent relaxation, the long component could potentially be relaxation processes of the dyes on the surface of the protein. We can conclude that solvent relaxation of Alexa488 can be measured using a time dependent energy loss of emission spectra yielding relaxation rates comparable to the expected ones.

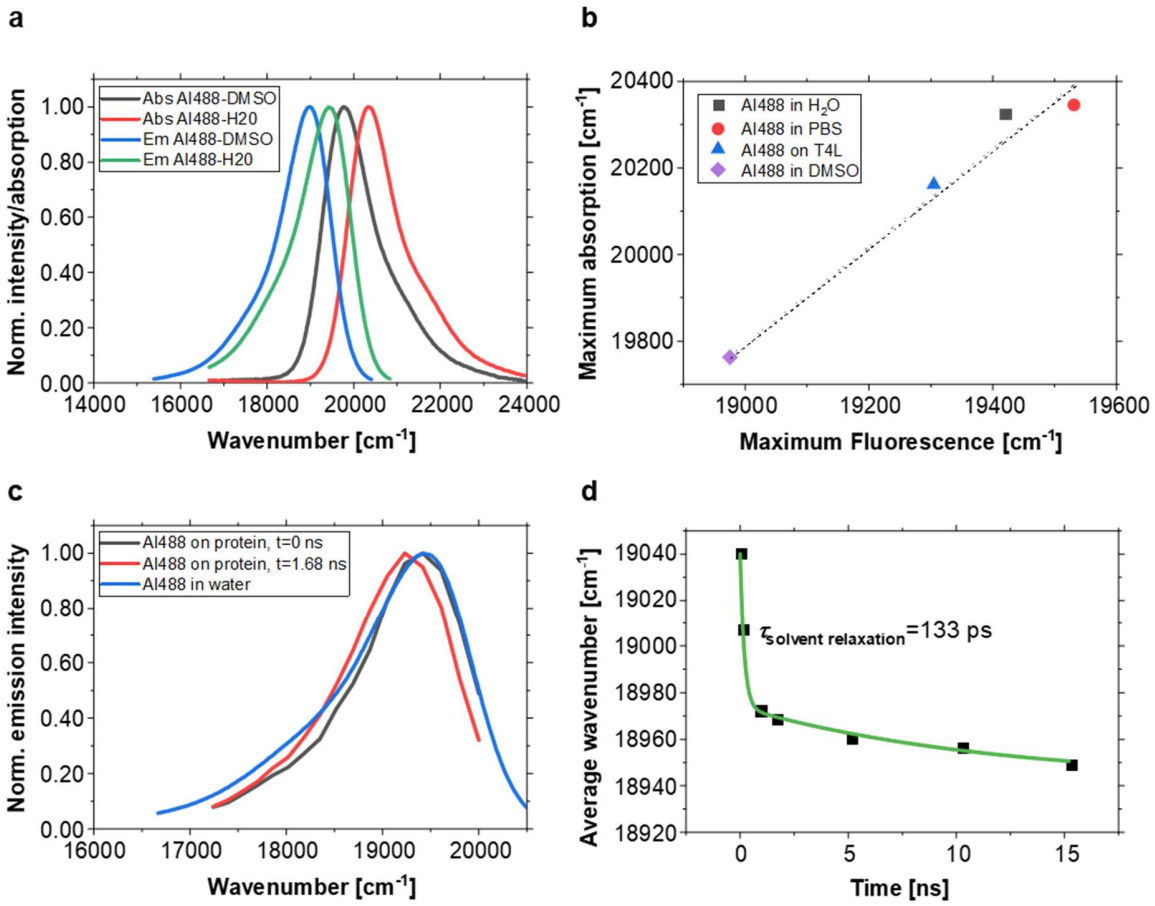


Figure 17: Solvatochromism of Alexa488 on hGBP1. a Emission and absorption spectra of Alexa488 in water and DMSO (data taken by Alexander Larbig). **b** Maximum fluorescence and absorption of Alexa488 in different solvents (data taken and analyzed by Alexander Larbig). **c** Red shift of Alexa488 linked to hGBP1 at amino acid 344. Spectra directly after excitation (black), after a waiting time of 1.8 ns (red) and in water (blue). **d** Energy decay due to solvent relaxation as a function of time. Green curve is based on a 2-exponential decay.

As a next step, this effect should be used to enhance accuracy in FRET derived distances. For this, the initial assumption that dyes away from the protein are more affected of solvatochromism was tested. Lifetimes were estimated using magic angle (ma) conditions $F_{ma}(t,\lambda)$. Species averaged lifetimes τ_x were fitted using in-house software (ChiSurf) described in literature [44] and freely available under https://github.com/Fluorescence-Tools/chisurf. One has to denote that for the used FRET sample a random labeling method was applied, meaning the donor/acceptor can be potentially on both available positions. The used test system was hGBP1 labeled at Q344C and V540C (see Figure 18) However, in the following it is assumed that the donor is at position Q344C, since the calculated solvent accessible surface area (SASA) is higher for Q344C compared to V540C (SASA(Q344C-C_β) = 0.04 Å^2, SASA(V540C-C_β) = 0 Å^2).

As a result, it was found that with an increase in the detection wavelength λ the lifetimes τ_x also increased (see Figure 19a). This effect can be seen for a free Alexa488 dye as well, but much smaller, since solvatochromism is faster. This goes in line with the assumption since dyes being in proximity to the protein can be quenched via static and dynamic quenching with specific amino acids, in which the dynamic quenching influences the measured lifetime. This phenomenon has

been investigated for instance for Alexa dyes in previous literature [44, 46]. Here, it was found that dyes are quenched by interactions with amino acids Trp, Tyr, Met and His via static and dynamic quenching. As it can be seen in Figure 18 these amino acids are accessible for the dye due to its flexible linker. As a further control, rotation correlation times were estimated. Here, $F_{||}$, $F_{\perp}$ and F_{ma} were fitted globally referring to methods described elsewhere [47]. Three rotational times are assumed, which are two for dye rotations and one for the slow, global movement of the protein itself, sensed by the attached dye. In figure 19b amplitudes of the rotation times relative to r_0 = 0.38 for dye position Q344C are shown as a function of λ. It can be seen, that with increasing λ, the amplitude of the global rotation decreases, indicating a higher flexibility of the dye. Also, the smallest lifetime τ_3, which is assumed to be the solvent relaxation, decreases in dependence to the detection wavelength (see Figure 19c). This might potentially occur since the dyes on the protein have a slower solvent relaxation compared to the dyes in free solvent. Furthermore, the amplitude of the highest lifetime x_1 increases, while the amplitude of the quenched lifetime x_2 decreases in dependence to the detection wavelength (see Figure 19d). That the amplitude x_3 assigned to the solvent relaxation decreases is to be expected since the effect of solvatochromism is declining due to the completion of the process. We can conclude that the initial assumption is confirmed and we can assume, that with an increase in the detection wavelength we mostly detect photons from dyes that are further away from the protein.

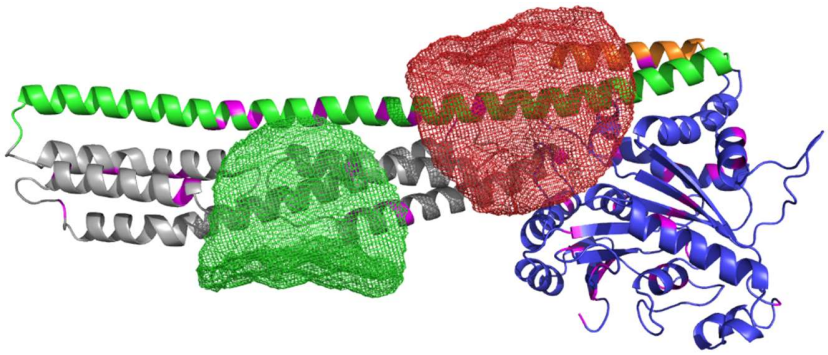


Figure 18: Labeled hGBP1 with quenching amino acids. Accessible volumes are shown for a very high fraction of trapped molecules. Quenching amino acids are shown in magenta. Dye position Q344 is shown in green, V540 in red.

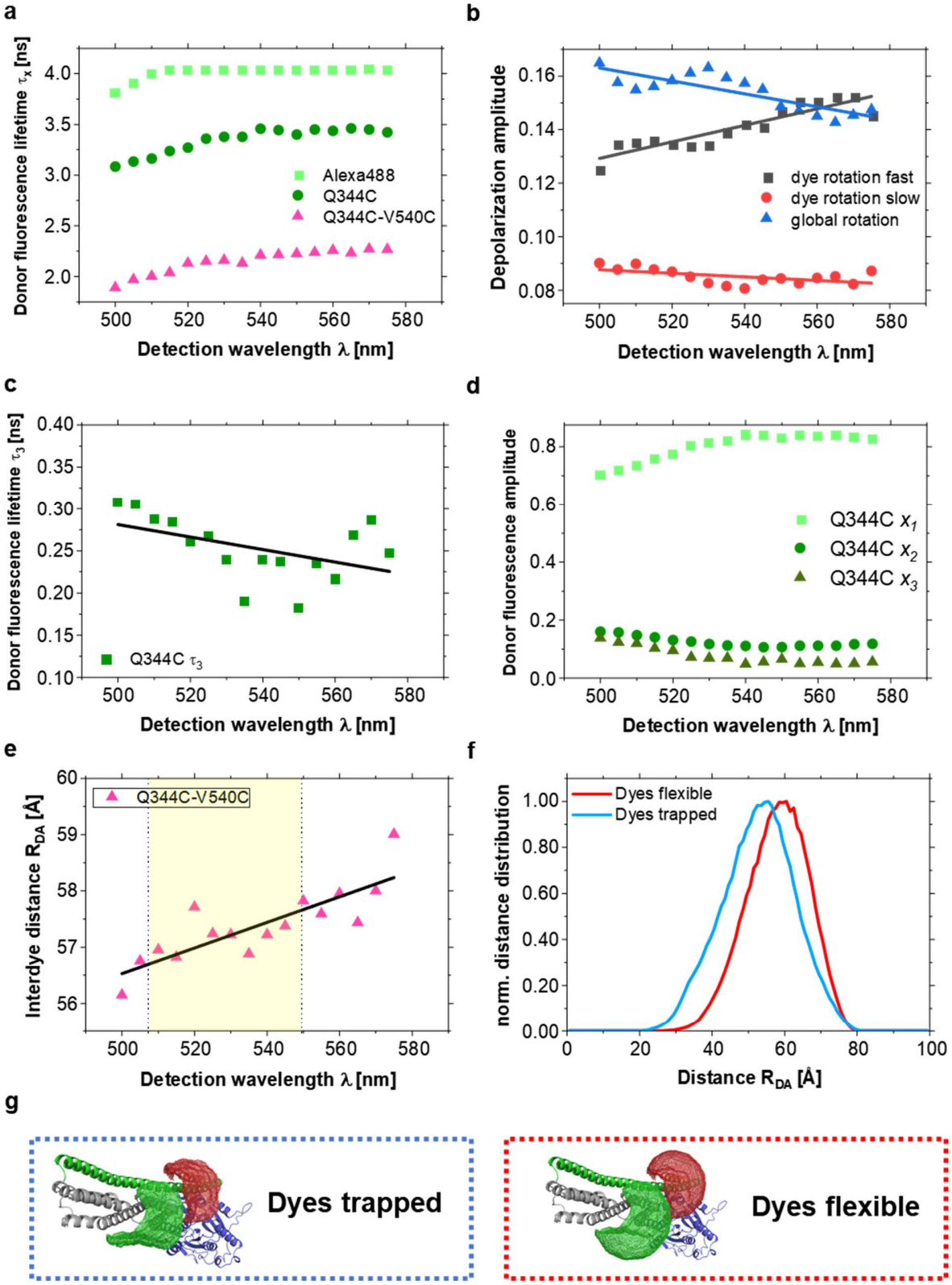


Figure 19: Solvatochromism of Alexa488 labeled to hGBP1. a Species averaged donor fluorescence lifetime τ_x as a function of the detection wavelength λ. Bright green data points are for Alexa488 free dyes in water and green data points for hGBP1 labeled with only a donor dye (Alexa488) at position Q344C.

Magenta points show the same for a double labeled sample with a donor and acceptor dye. **b** Amplitudes of rotation correlation times relative to fundamental anisotropy r_0 of the donor only labeled sample Q344C. Black points indicate the amplitude of the fast dye rotation with $\boldsymbol{\rho_{dye,fast}}$ = 0.2 ns, red points of the slow dye rotation $\boldsymbol{\rho_{dye,slow}}$ = 1.5 ns and blue of the global rotation of the protein $\boldsymbol{\rho_{protein}}$ = 35 ns. Permanent lines are based on a linear fit indicating the tendency. **c** Shortest lifetime $\boldsymbol{\tau_3}$ of the Q344C donor only sample as a function of the detection wavelength. Straight line indicates a decreasing tendency. **d** Amplitudes of the multi exponential fit of the the Q344C donor only sample. **e** Interdye distances $\boldsymbol{R_{DA}}$ based on the species averaged lifetimes $\boldsymbol{\tau_x}$ for the measured FRET sample Q344C-V540C as a function of the detection wavelength $\boldsymbol{\lambda}$. Typical detection window analyzed in greater detail in chapter 3.3 shown in bright yellow. **f** Distance distributions for the diffusive dye behavior for the two assumed cases when the dye is trapped on the surface of the protein (blue) or only flexible (red). In both cases simulations were performed using a linker length of 20 Å, a linker width of 2 Å and a dye radius of 3.5 Å. For trapped dye scenario a trapped fraction of 99 % was taken, for free dye scenario 0 %. **g** Accessible volumes on the positions Q344C and V540C of the hGBP1 for only trapped dyes (blue) or flexible dyes (red).

Finally, after testing the occurrence of solvatochromism and the confirmation of the assumption, this effect is applied to study FRET derived distances. Based on measured lifetimes τ_x, interdye distances R_{DA} are calculated. A constant Förster radius of R_0 = 52 Å was assumed. As a result, it was found that the measured distances increased with the detection wavelength (see Figure 19e). These distances can be compared to expected distances based on accessible volumes and 3D structures of the hGBP1 (which will be derived in chapter 4.3) using methods described in literature [48] (see Figure 19f). As an outcome, for the major state of the hGBP1 one expects for only flexible dyes (see Figure 19f) an interdye distance of R_{DA} = 58 Å, whereas only trapped dyes would yield to R_{DA} = 53 Å, which shows the same tendency for the measured value of low detection wavelength $R_{DA}(\lambda\ low)$ = 56 Å and $R_{DA}(\lambda\ high)$ = 59 Å. The minor state of the hGBP1 is in that case neglected since only a single distance was fitted but would yield to a similar distance change $R_{DA}(free\ dye)$ = 59 Å and $R_{DA}(trapped\ dye)$ = 53 Å). Concluding, this shows that for this specific dye pair an effect of the fraction of flexible and trapped dye predicted by calculated interdye distances could be measured using solvatochromism.

3.4.3 Conclusion

The effect of solvent relaxation on the system dye-solvent can be measured using detection wavelength dependent measurements. An energy loss was observed on a time scale relevant for solvent relaxation. The assumption of the average dye position being close to the protein for low detection wavelength or away from it for high is confirmed using two experimental observables, which are the lifetime and the anisotropy. This yields to a measurable effect on the interdye distance. Therefore, it would be advantageous using the detection wavelength as an additional parameter in the future, suggesting an $R_{DA}(detection\ wavelength)$. In case of measuring at high energy values of the photons, one runs into the trouble that the majority of them are emitted by dyes being close or trapped to the protein, which one wants to avoid. On the other hand, measuring photons with a high red shift, one avoids the trapped dyes but loses signal since the maximum of the emission spectra is rather blue then red. Therefore, it would be advisable to avoid only too small detection wavelengths. For this specific dye pair, label position and protein, measuring at wavelengths between 510 and 550 nm would be advisable assuring a higher accuracy in the measured distances since the impact of trapped fraction will be reduced. This window is what our laboratory typically uses in its single molecule MFD studies (see chapter 3.3). This approach can potentially solve problems with dye pairs that show very high combined residual anisotropies like in the case of some dye pairs used in the global FRET study in chapter 3.2. Here, using the elaborated experimental advices will reduce the anisotropy values allowing these measurements to be analyzed accordingly instead of being faulted.

3.5 Conclusion of chapter 3

In the first section presenting the global FRET study, it could be seen that the accuracy obtained using FRET derived distances and obtained dynamical models of proteins show very good agreement with setups and analysis methods from all around the world. While the MalE protein was used to study distances, U2AF65 was a challenging system having a complex dynamic behavior. Nevertheless, good agreement in the obtained relaxation times of the conformational exchanges was achieved. However, problems occurred due to a user bias in the analysis which is mainly caused by different approaches in estimating the correction factors. We answered this with the introduction of a generally applicable workflow to estimate calibration factors in the second section. The elaborated workflow was tested within a blind study at the AG Seidel and very good agreement with minimum user bias was found. On top we presented a new, unique approach in estimating the γ-factor by using a set of free dyes and measurement of the installed optical components in the setup. Doing so, one entangles the calibration process with the actual measurement of the samples, ensuring a high reliability of the method. Another problem from the global FRET study was answered in the third section. Here, I suggest that using the detection wavelength as an additional parameter, one can avoid a high fraction of trapped dyes. This is achieved via avoiding small detection wavelengths. This approach might not be generally applicable since the polarity of dyes differ and the effect could become minor. However, we tested this for the commonly used dye pair Alexa488 and Alexa647 and found a measurable effect.

4 Applications – smFRET studies resolving the dynamics of biomolecules under manifold settings

4.1 Overview

In the previous chapter it could be seen that smFRET studies can measure structures and monitor dynamics of biomolecules at a high precision. As so, the following chapter presents three different biomolecules which I studied using different fluorescence spectroscopy techniques and methods. Here, the main focus will be the influence of different surrounding conditions on biomolecular dynamics.

As a first, a nucleic acid which is showing fast conformational exchanges under the addition of magnesium ions (chapter 4.2) is studied. To do so, a new experimental approach is introduced. This is a device that enhances obtained signals and furthermore enlarges the observation time of an object measured using confocal single-molecule spectroscopy. As development and manufacturing is done by L. Morales-Inostroza, Stephan Götzinger and Vahid Sandoghdar, I tested and applied it to a biomolecular system, namely the Holliday Junction. Here we could take measurements that can follow the conformational changes of the HJ in “real time”.

Next, a human protein, responsible for innate immune responses within the cell, is studied. Here, we observe that a small change which is adding a farnesyl moiety in the protein modifies its dynamic behavior. This protein exhibits extensive dynamical behavior, which could be properly characterized by combining old and new data for an improved model of the protein. Furthermore, the protein was investigated with its natural posttranslational modification, the farnesylation. This relatively minor change to the protein lead to major changes in the dynamical behavior, which gave insight to the biological relevance of this modification.

Finally, a project investigating a toxin is presented. In this project, the molecule displays a complex dynamical behavior which is only induced by a pH shift of the solvent. The molecule is then active on different time scales which can be modulated via mutations of the protein. I use remarkably different experimental approaches to resolve a global picture of the fundamental mechanism of the molecule.

4.2 Resolving the energy landscape of a Holliday Junction using an optical antenna

Luis Morales-Inostroza,[1,2,3,*] Julian Folz,[4,*] Ralf Kühnemuth,[4] Suren Felekyan,[4] Franz-Ferdinand Wieser,[1,2,3] Claus A.M. Seidel,[4,†] Stephan Götzinger,[3,1,5] and Vahid Sandoghdar[1,3,‡]

[1] *Max Planck Institute for the Science of Light, 91058 Erlangen, Germany*

[2] *Max-Planck-Zentrum für Physik und Medizin, 91058 Erlangen, Germany*

[3] *Department of Physics, Friedrich-Alexander-Universität of Erlangen-Nürnberg, 91058 Erlangen, Germany*

[4] *Chair for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

[5] *Graduate School in Advanced Optical Technologies (SAOT), FAU, D-91052 Erlangen, Germany*

[*] *These two authors contributed equally*

[†] *Correspondence email address: cseidel@hhu.de*

[‡] *Correspondence email address: vahid.sandoghdar@mpl.mpg.de*

Manuscript stage: Close to submission, final reading.

Current version of the manuscript attached under chapter 7.2.

Preliminary author contributions: LMI and JF wrote the initial draft of the manuscript. LMI manufactured the OFA and performed basic tests on it at the MPI Erlangen. JF, SF, RK and LMI performed sm-fluorescence measurements at the HHU Düsseldorf. JF, SF, LMI and RK analyzed and interpreted the data. JF managed the samples used at HHU Düsseldorf. LMI and FW performed simulations explaining the working principle of the OFA. CAMS, SG and VS designed the research and supervised the project.

4.2.1 Background: Molecule of interest

In this study, I mainly investigated the Holliday Junction itself, while I used free dyes to first test and benchmark an optical antenna enhancing the obtained signal. The Holliday Junction (HJ) is a nucleic acid structure and consists of four double stranded DNA branches forming a DNA 4-way junction. It is named after R. Holliday, who proposed it in 1964 [49]. In this study, we used following sequence for the four branches α, β, γ and δ: α(5´-CCT AAT TAC CAG TCC AGA TTA ATC AGT ACG), β(5´-CGT ACT GAT TAA TCT CCG CAA ATG TGA ACG), γ(5´-CGT TCA CAT TTG CGG TCT TCT ATC TCC ACG) and δ(5´-CGT GGA GAT AGA AGA GGA CTG GTA ATT AGG). For FRET experiments the α and β strands were labeled using Alexa488 as donor and Atto647N as acceptor dye: α (5´-CCT AAT T(Alexa488)AC CAG TCC AGA TTA ATC AGT ACG), β (5´-CGT ACT GAT T(Atto647N)AA TCT CCG CAA ATG TGA ACG). The HJ is formed during DNA strand exchange of homologous recombination and is fundamental in genetic recombination [50, 51]. Under addition of Mg^{2+} ions the HJ forms two folded structures fluctuating between each other, where the branches are stacked in a parallel or anti parallel fashion (see Figure 20) [52]. The

formation of folded structures is due to the metal ions shielding the electrostatic repulsion of the junctions [53]. Therefore, dynamic exchange between these two states depends on the amount of metal ions in the buffer [54]. It was found [55, 56] in experiments with a varying concentration of Mg^{2+} ions that the HJ experiences a complex four state kinetic exchange, where two of the states have a Mg^{2+} ion bound to it, resulting in static states. Additionally, the other two states show a similar structure as the two static ones but on top displaying dynamic small-scale fluctuations. However, in this work, we use the simplified two state model of the HJ based on the Mg^{2+} concentration shifts the dynamic equilibrium between the two states. To go from one folded structure to the other, the molecule must pass a cross shaped structure. It is an ongoing discussion whether this open cross state is a stable intermediate state or a short-lived transition state [54, 57, 58].

4.2.2 Aim of the study

Confocal fluorescence spectroscopy of freely diffusing molecules faces two eminent limitations, the major one being the limitation of the number of photons one can get while the molecule is in the focus. The minor one is the maximum observation time which is defined by the translational diffusion of the molecule passing through the confocal volume. To overcome these, one can use for instance optical antennas [59] enhancing the obtained fluorescence signal and experimental approaches that can trap the molecule (for instance the “Abel trap” [60]). In this study, we improved the performance with a single modification that addressed both issues. The optical antenna was designed and manufactured in the working group of Vahid Sandoghdar. To do so, we combined it with the single-molecule Multiparameter Fluorescence Detection (sm-MFD) approach established in our laboratory. After initial manufacturing and testing was performed in the working group Prof. Sandoghdar, we installed the OptoFluidic Antenna (OFA) in our laboratory for a more detailed characterization via measuring fluorescence-based parameters using a free dye as a convenient test system. Finally, we wanted to apply the technique to the HJ taking advantage of the enhancement factor in the brightness. To do so, I first tested the behavior of the HJ in the antenna to ensure that it causes no disruptive interaction. Subsequently, I designed an experiment using high excitation rates, resulting in intensities high enough to measure the transition of the HJ from one stacked conformation to the other (see Figure 20). Doing so, the HJ has to pass an open, cross shape state. In this experiment, I aimed to achieve a signal high enough to study the existence of the cross-shape state as a possibly fast transient state at time resolution on which we can track the conformational changes in “real time”. To do so, I wanted to follow the FRET efficiency trajectory without any deeper analysis on a biological relevant time scale.

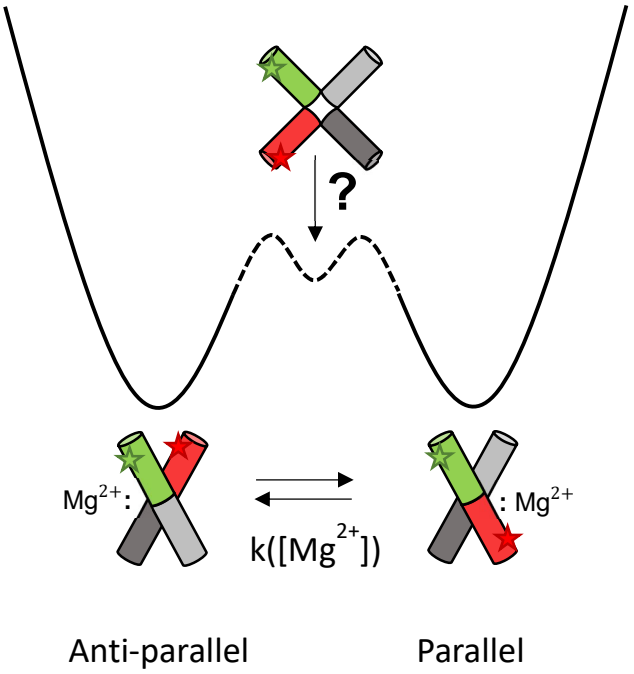


Figure 20: Energy landscape and sketches of the Holliday Junction. Upon addition of Mg^{2+} the HJ fluctuates between two folded structures with a potential cross shape structure in between. Donor labeled branch of the HJ is shown in green, acceptor labeled branch in red.

4.2.3 Results

Benchmarking the OFA

The OFA was invented and manufactured by L.M. Inostroza, S. Götzinger and V. Sandoghdar within the scope of a dissertation [61]. It is based on a simple glass pipette which is positioned above the confocal volume of an inverted microscope (see Figure 21). Thereby an additional water-air interface is formed, closely above the confocal volume, leading to a total reflection of the emitted photons. These photons will be back reflected into the objective, leading to a higher obtained signal. In order to test its performance, experiments were performed in the laboratories of the AG Seidel. There, I used a simple excitation power series using the free dye Rhodamin110 (see Figure 22) to measure the brightness enhancement. The brightness is defined as the ratio of the obtained signal divided by the number of molecules in focus being in a bright state (triplet corrected). Via Fluorescence Correlation Spectroscopy (FCS) one can obtain these numbers fitting the correlation curve $G(t_c)$ that includes the triplet state and translational diffusion in a 3-dimensional Gaussian shaped volume:

$$G(t_c) = b_0 + \frac{1}{N_{bright}}\left[1+\frac{t_c}{t_d}\right]^{-1}\left[1+\frac{t_c}{t_d\left(\frac{z_0}{\omega_0}\right)^2}\right]^{-1/2}\left[1-|A|+|A|e^{t_c/t_A}\right] \qquad (4.2\text{-}1)$$

where b_0 denotes the offset, N_{bright} the number of molecules in a bright state in focus, t_d the translational diffusion time, A the amplitude for the triplet state and t_A the time constant for the triplet state. To estimate the brightness Q of a molecule it follows then with the fluorescence signal of the donor after donor excitation:

$$Q = \frac{F_{D|D}}{N} = \frac{F_{D|D}}{N_{bright}}(1 - A). \quad (4.2\text{-}2)$$

The data points for the brightness follow a saturation curve, which is defined as

$$Q(I_0) = Q_{sat}(1 - e^{-\frac{I_0}{I_\sigma}}) \quad (4.2\text{-}3)$$

with the excitation power I_0 the cross-section excitation power I_σ and the saturated brightness values Q_{sat}. This model fitted a global I_σ value of 64 µW. I found that the brightness of a fluorescent molecule increased when the antenna is used (see Figure 22a,b), depending on the antenna's height with respect to the glass surface. It was also observed that the translational diffusion time of the molecule increased significantly (see Figure 22c). This effect is assumed to be caused by the water-air interface and quantitatively reproduced by L.M. Inostroza using simulations, which can be found in the manuscript (see 7.2). The height of the OFA was measured using a CCD camera in combination with a simple torch creating a shadow of the OFA on the glass surface leading to a rough estimate, which is afterwards specified by the piezo-position readout (see Figure 22d). Here, a linear dependence on the brightness enhancement was observed.

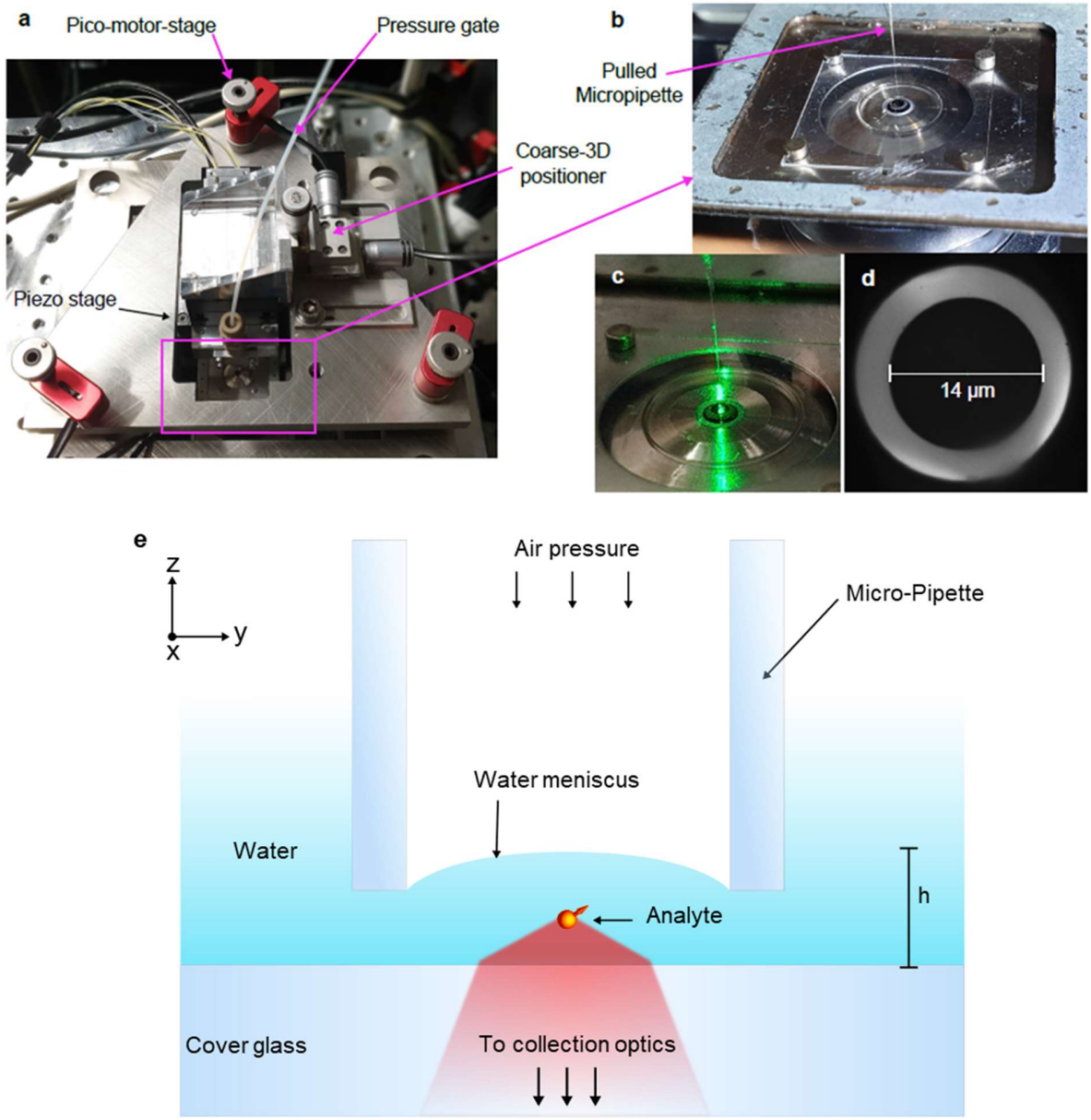


Figure 21: Setup of the optofluidic antenna. a AFM-like plate to position and align the OFA. Alignment of the plate is done via a pico-motor stage, fine alignment and control of the antenna via a piezo-stage controlled with a closed loop piezo. **b** Pulled micropipette in position above the objective. **c** Laser light hitting the OFA to ensure a close position to the focus. **d** Image of micropipette ending with a camera showing a typical diameter. Figure and part of the caption taken from [61]. **e** Sketch of a single molecule (analyte) diffusing through the OFA. The thickness of the water layer is indicated with h = 500 nm.

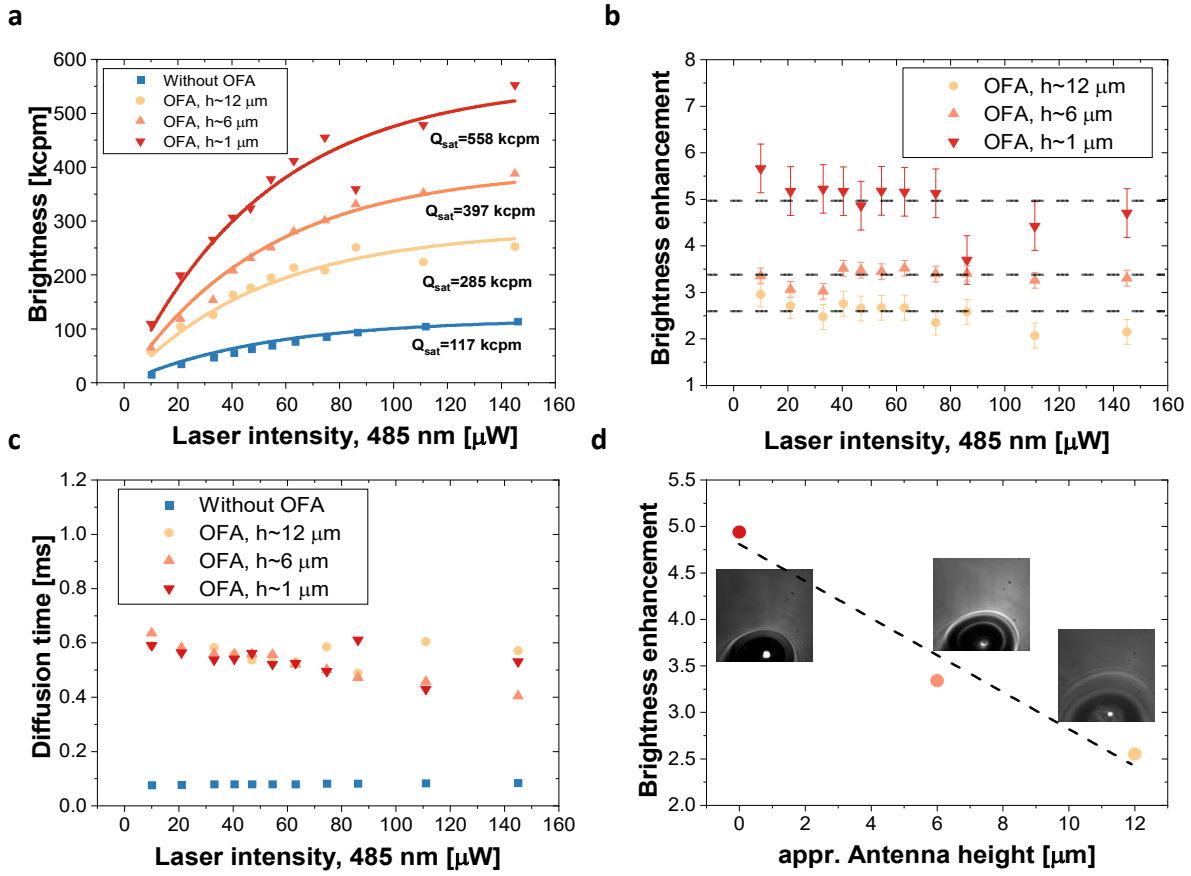


Figure 22: Rhodamin110 Power Series. a Power-dependent brightness of confocal measurements of freely diffusing Rhodamin110-dye (Rh110) obtained via fluorescence correlation spectroscopy (FCS) using the antenna at different height-positions h and without antenna as reference. Each data point represents a 10-minute measurement where the normalized autocorrelation ($G(t_c)$) function was fitted using an equation with two anti-bunching terms for triplet and diffusion and a 3-dimensional Gaussian shape confocal volume (see (4.2-1). Brightness values are following a saturation curve (see eq. (**4.2-3**) **b** Brightness enhancement of every data point using the antenna compared to the brightness-values without the antenna. Dashed gray line indicates the average brightness enhancement for a specific antenna height position. **c** Analysis of the translational diffusion time of Rh110 using the antenna obtained from AC function analysis. Using the antenna results in higher translational diffusion times compared to without. **d** Average brightness enhancement shown for each height position of the antenna. Pictures next to data points are taken with white light illumination from the top and a CCD camera. The laser focus is identified as the white spot in the middle of the circular shaped antenna. The antenna position and height were calibrated via live imaging where the difference between the “white shadow” of the antenna on the surface of the cover glass and the antenna itself was used to set an approximate distance, which was afterwards specified from the piezo-position readout.

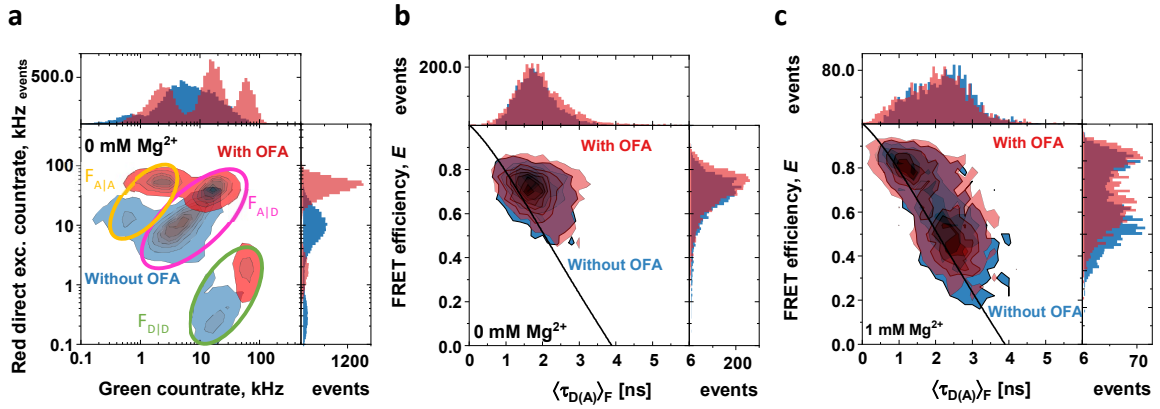


Figure 23 Benchmarking the OFA using smMFD. Blue shows the histograms of the measurements without OFA, red of the ones with OFA. **a** 2D smMFD measurement of the labeled HJ using Alexa488 as donor and Atto647N as acceptor dye. On *x*-axis the green countrate is shown, on *y*-axis the red countrate after acceptor excitation. Three populations are identifiable, where the one with high green and low red countrate and is showing the molecules labeled with only a donor molecule, the one with high red and low greencountrate the molecules labeled with only an acceptor molecule and the population in the middle the double labeled molecules. Yellow circle indicates the fluorescence emitted from the acceptor under acceptor excitation $F_{A|A}$, magenta circle the fluorescence emitted from the acceptor under donor excitation $F_{A|D}$ and the green circle the fluorescence emitted from the donor under excitation of the donor dye $F_{D|D}$. **b** Measurement of the HJ under 0 mM Mg^{2+} conditions. Plotted are the donor fluorescence lifetime $\langle \tau_{D(A)} \rangle_f$ against the FRET efficiency *E*. Black line indicates the static FRET line. **c** Shows the same for the measurement under 1 mM Mg^{2+} conditions.

As a next step, the performance of the antenna under single-molecule conditions was tested (see Figure 23). Here, the results from the free dye power series were confirmed and enhanced signals could be observed as well, independent of the spectral range (see Figure 23a). While the signals increased by approximately a factor of 4-5, an important condition to use the OFA is that it does not perturb the biomolecules in any way, meaning that structure and dynamical behavior are unaffected. This was tested via measuring FRET derived distances and dynamics of the HJ under $0\ \mathrm{mM}$ (Figure 23b) and 1 mM Mg^{2+} (Figure 23c) conditions. The measurements with OFA show no fundamental difference to the ones without OFA. Further analysis and testing of the difference in the dynamical behavior was done using fFCS (filtered Fluorescence Correlation Spectroscopy) and can be found in the manuscript (see 7.2). In summary, the OFA is able to enhance the signals significantly and additionally slows down the translational diffusion leading to longer observation times, without perturbing the biomolecule. This feature was used to study the HJ in greater detail.

Resolving the dynamics of the HJ using the OFA

As worked out in the previous section, the OFA increases both the observation time in which the molecule can be observed and in addition, the overall signal. Both effects are very useful and were tested using the Mg^{2+} dependent dynamics of the HJ.

As introduced in 4.2.1, the exchange of the HJ is fast at low concentrations of Mg^{2+}, and slows down at higher concentrations. Slow dynamics are challenging in smMFD because they are difficult to track while the molecule is in the focus. As an example, a kinetic with relaxation times of $10\ \mathrm{ms}$ will be rarely observed for a molecule that diffuses through the focus within 1 ms. In case of the HJ we used 1 mM Mg^{2+} leading to a relaxation time of approximately 600 µs, obtained by a Photon Distribution Analysis (PDA). Without the OFA, an average translational diffusion time of ~0.7 ms is observed, whereas with the application of the antenna we measure ~2 ms. To analyze the dynamical behavior a PDA was performed (see Figure 24). As an outcome, using the OFA leads to a larger translational diffusion causing a longer time window where one can observe the molecule. Hence there is a higher probability of a transition event within the confocal volume, where the molecule goes from one state to the other, leading to dynamical averaging of the experimental observable. This can be seen by the fact that the dynamic fraction with an average FRET efficiency peaked at $E = 0.7$ in the PDA without OFA (see Figure 24a) is much less populated in comparison to the one with OFA (see Figure 24b). Therefore, it can be summarized that the trap feature of the OFA allows for experiments with higher observation times resulting in an increased accuracy due to higher statistics in analyzing dynamics on the time scale of the translational diffusion.

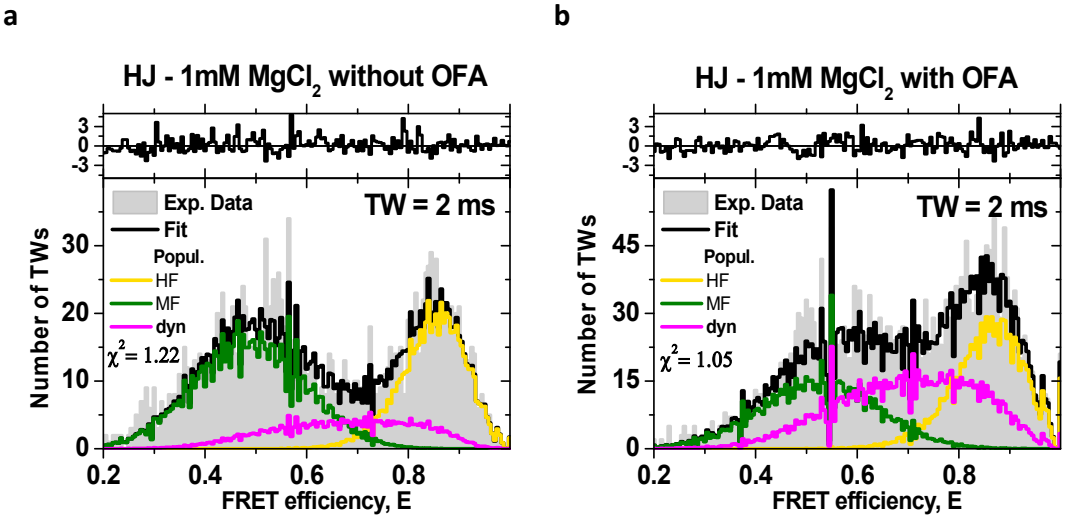


Figure 24: Resolving slow dynamics of the HJ using the OFA. Photon Distribution Analysis (PDA) of HJ under single-molecule conditions using 1 mM Mg^{2+} without (**a**) and with (**b**) OFA. Medium FRET efficiency (MF) population is shown in green, High FRET efficiency (HF) population is shown in yellow. Dynamic population in between these two states is shown in magenta. Figures are shown for a Time Window (TW) size of 2 ms.

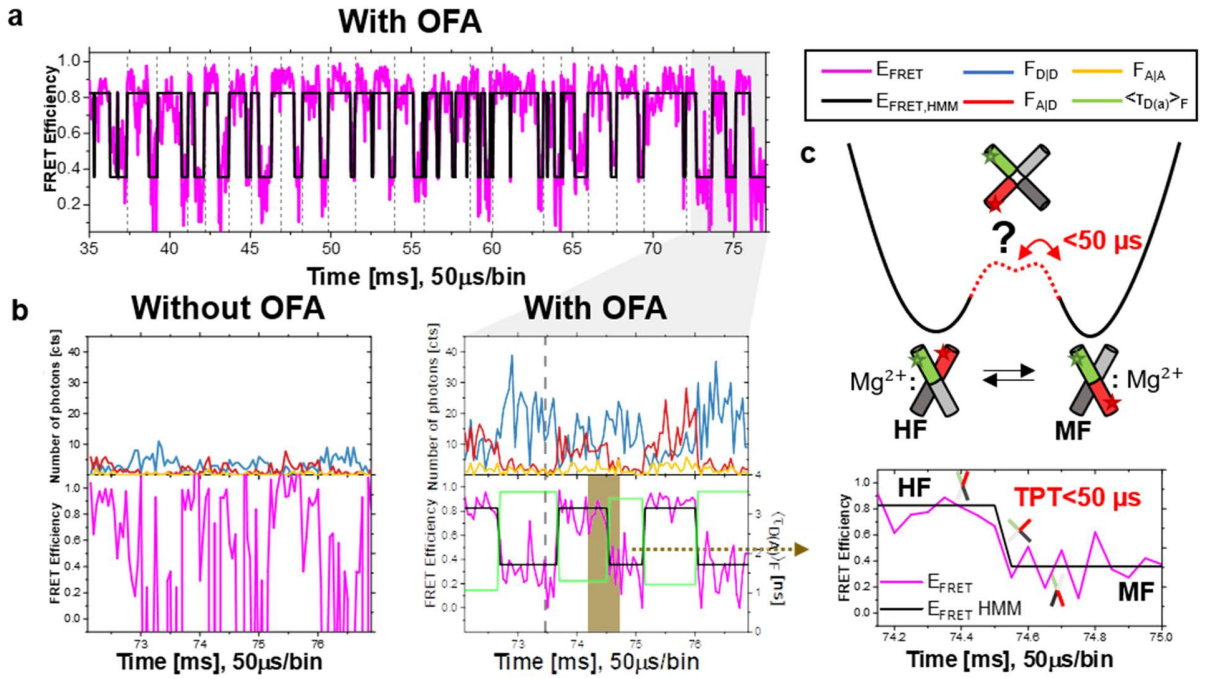


Figure 25: HJ single-molecule burst trajectory. FRET efficiency trajectory consisting out of individual bursts merged together using a time binning of 50 µs and a buffer containing 0.2 mM Mg^{2+}. **a** Dashed grey line indicates start and endpoint of individual bursts. Magenta line represents the measured FRET efficiency calculated from corrected green and red signal, black line shows the fitted levels using Hidden Markov Modeling (HMM). **b** Trajectory of the measured signals using the OFA (right) and as comparison the same trajectory without OFA with an assumption of an enhancement factor of 4 (left). Shown is the donor signal under donor excitation $\boldsymbol{F_{D|D}}$ in blue, acceptor signal under donor excitation $\boldsymbol{F_{A|D}}$ in red, and acceptor signal under acceptor excitation $\boldsymbol{F_{A|A}}$ in yellow. Additionally, the fluorescence donor lifetime $\langle \boldsymbol{\tau_{D(A)}} \rangle_f$ is shown in green for the identified dwell times. **c** Resolved energy landscape of the HJ with the two folded structures with a Mg^{2+} ion bound. FRET efficiency trajectory shows fast transition path times (TPT) from one state to another indicating an only very small energy barrier for the cross-shape state of the HJ.

To use the signal enhancement of the OFA high excitation intensitiy (I_{exc} = 800 µW) experiments were performed in order to obtain maximum signal for studying the transitions in real time. The experiment shows signals which are exceedingly high. Using standard conditions with normal excitation rates (e.g., I_{exc} = 80 µW), one typically obtains average countrates of 50 kHz. Using the OFA with high excitation intensities (I_{exc} = 800 µW) we measured average count rates of around 700 kHz, with peak values of up to 1 MHz. Using this signal, we can assemble a FRET efficiency trace out of individual single-molecule events (see Figure 25). Now, this exciting experiment not only gives information about the dwell times the molecule stays in one state until it goes to another, but also gives indications about the transition time. The transition time or transition path time is the time the molecule takes to go from one structure to another. Using the software “HaMMy” [62] a Hidden Markov Modelling (HMM) algorithm was applied to estimate the FRET efficiency values and states in the trajectory and it was found that the molecule transitions from one folded structure to another within a single bin (bin time is 50 µs), while no significant populated intermediate state was found. Therefore, the transition time of going from one stacked state (e.g., HF) to the other, is much faster than the bin time of 50 µs. This suggests that the cross-shape state of the HJ which it has to pass to go from one stacked state to the other is a fast decaying transition state with a dwell time less than the used bin time of 50 µs.

4.2.4 Conclusion

The OFA is a simple device that can be comparably easily installed and calibrated on already existing optical setups based on an inverted microscope without further changes. It was tested that it enhances the signal by a typical factor of 4-5, and enlarges the observation time of a biomolecule significantly by slowing down its translational diffusion. Also, it fulfills its basic condition which is that it does not perturb the biomolecule in any way and also does not show any spectral dependence so that all existing analysis techniques can be applied without further considerations. This makes the OFA advantageous over experimentally more challenging approaches like the ABEL trap [60, 63, 64], plasmonic-nanoantennas [65, 66] or DNA origami nanoantennas [67].

In our laboratory the performance and usability of the OFA was tested and used to study the dynamical behavior of the HJ. With the OFA it was possible to follow the behavior of the HJ in “real time”, meaning that the dwell time of one state was tracked and can be followed by a FRET efficiency trajectory, until it goes over a transition to the other state. As using these high signals, the transition path time could be tracked down to less than 50 µs, which is the binning time of the trajectory. These types of experiments could be so far only done with higher binning times leading to a lower resolution time [68-70]. Also, they are performed under near physiological conditions, so that they can be very informative to study the dynamical behavior of biomolecules in “real time”, as proposed in literature [1].

4.3 Control of dynamics and function by a single farnesylation of hGBP1

Julian Folz[1,#], Paul Lauterjung[2,#], Christian Herrmann[*], Claus Seidel[*]

Author list preliminary.

[1] *Chair of Molecular Physical Chemistry, Heinrich-Heine University Düsseldorf*

[2] *Physical Chemistry I, Faculty of Chemistry and Biochemistry, Ruhr-University Bochum*

[#] *These authors contributed equally*

[*] *Corresponding author*

Manuscript stage: Manuscript in final stage of preparation.

Current version of the manuscript attached under chapter 7.3.

Author contributions: JF and PL wrote the initial draft of the manuscript. PL prepared the samples and measurement conditions. JF performed sm-confocal fluorescence experiments and analyzed data. PL performed confocal fluorescence imaging spectroscopy and analyzed data. JF and PL interpreted data. CH and CAMS designed the research and supervised the project. JF, PL, CH and CAMS interpreted the data and wrote the manuscript.

4.3.1 Background: Molecule of interest

The human Guanylate Binding Protein 1 (hGBP1) is one of seven GBP paralogues within the human body belonging to the dynamin superfamily of large GTPases, more specific to the class of interferon-γ induced effector molecules. Hence, it has its key roles in the human cell's immune response to a huge variety of threats [71], which includes defense against bacterial, viral and protozoan attacks [72-74]. Even antitumor activities are reported [75]. It can be found in both the cytoplasm and on membranes in the cell. A crucial feature of the hGBP1 as a large GTPase is its ability to bind a GTP and to hydrolyze it to GDP and GMP gaining its biological functions [75].

The structure of the hGBP1 (see Figure 26) is divided into different domains, which is the large GTPase domain (LGD, aa 1-308, shown in blue), the middle domain (MD, aa 309-480, shown in grey) and the GTPase effector domain (GED, aa 481-592, shown in green and orange). On top, the hGBP1 possesses a farnesyl moiety (shown in magenta) after the GED, covalently linked to C589, resulting in the farnesylated state $hGBP1_{farn}$, or in case of absence, the non-farnesylated state $hGBP1_{non\text{-}farn}$. The farnesyl moiety is able to interact with a hydrophobic binding pocket positioned at the α9-helix (involved aa: H378, Q381, K382, A385) and the α12-helix (involved aa: Y524, H527, L528, L531) [76] or, at a later stage binding to a membrane. The function of the protein is connected to hydrolysis and its pathway going from a monomeric to an oligomeric state [77]. While the non-farnesylated hGBP1 can only dimerize [78], the farnesylated hGBP1 can also oligomerize, indicating a significant difference between the farnesylated and non-farnesylated hGBP1. However, to perform this pathway a nucleotide or an analogue is needed. While its natural nucleotide is GTP, whose hydrolyzation was studied with the LGD as the catalytic center of activity [79], in this work, the GTP analogue GDP-AlF_x was mainly used, simulating the transition state between GTP and GDP. With this analogue, oligomerization was observed before [77, 80, 81], but details about structural and dynamical behavior performing its function remain unresolved.

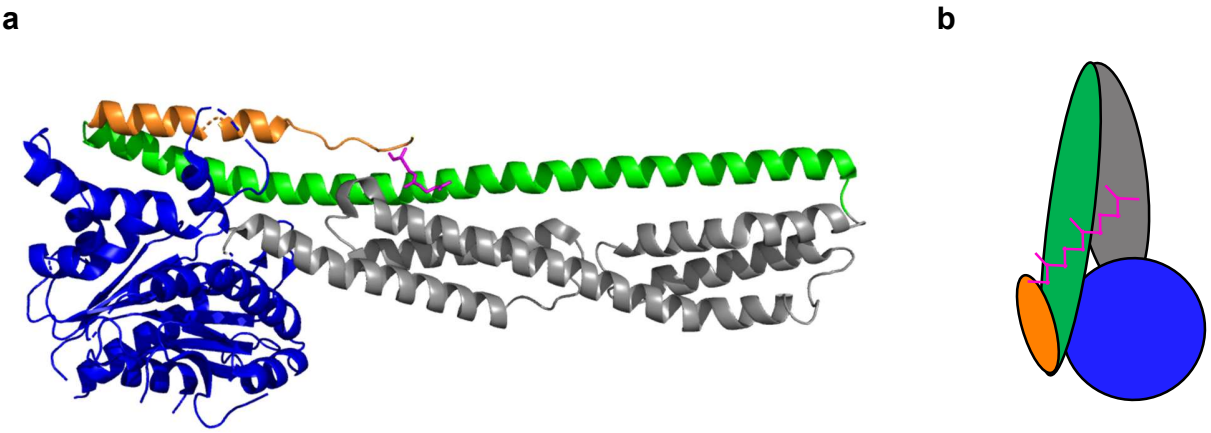


Figure 26: Structure of the farnesylated hGBP1. LGD is shown in blue, MD in grey, GED is divided into α12-helix shown in green and α13-helix shown in orange. The farnesyl moiety is shown in magenta. Shown is the nucleotide free form. **a** 3D-structure of $hGBP1_{farn}$ (pdb 6k1z). **b** Schematic representation of $hGBP1_{farn}$.

4.3.2 Aim of the study

As introduced in the previous section the hGBP1 plays a crucial role in the immune response of the human body. Mechanism, structure and dynamical behavior of the hGBP1 have to be understood since it is the basis of every protein's function [82, 83]. In a previous study [84] the hGBP1 structure and dynamic of the non-farnesylated state were measured using a joint study of EPR, FRET and SAXS data combined with structural modeling. It was shown that the $hGBP1_{non-farn}$ possesses two different monomeric states, M_1 and M_2 under ambient conditions. These two states are in a fast-conformational dynamic exchange (2-300 µs). The main difference of these two states is a flip of the α12-helix around the MD of the protein, resulting in a "mirrored" conformation. This behavior becomes meaningful when the hGBP1 is dimerizing. The dimer is initially formed via a contact of the LG of two monomers, resulting in the LG:LG dimer [85, 86]. In case of a homo dimer consisting out of two M_1 monomers, M_1: M_1, the α12/13-helices are on different sides in respect to each other. It changes in case of a hetero dimer M_1: M_2, enabling the molecule to connect the α12/13-helices. This dimer is more stable and potentially the relevant one considering the next step of the protein, which is the polymerization. However, the non-farnesylated hGBP1 can only form dimers, but no higher-order oligomers. In this study we build on the previous study and extend the previously used FRET network consisting of 12 FRET pairs with 11 new FRET pairs in order to refine the structural models for the M_1 and M_2 states of the hGBP1. By doing this, we want to resolve the structures of each conformer in greater detail. As a next step, we extend the study to the farnesylated hGBP1 and resolve the similarities and differences in the structural and dynamical behavior in comparison to the non-farnesylated hGBP1. To do so, five new mutants are labeled, connecting each domain of the hGBP1. Via adding the substrate GDP-ALF_x and a varying amount of wild type hGBP1, we study the whole pathway of the $hGBP1_{farn}$ from monomeric to its oligomeric state. After that, we can carefully compare the results from the non-farnesylated hGBP1 to the farnesylated one in order to understand the differences and therefore the function of the physiologically relevant $hGBP1_{farn}$.

4.3.3 Key Results

Structural refinement of hGBP1

Structural refinement of the M_1 and M_2 states of the hGBP1 is done via incorporating the new FRET pairs to the existing data which is a combination of EPR, FRET and SAXS. To find the most informative new FRET pairs, an algorithm for FRET based structural modeling is used [17]. After the selection, new mutants were expressed, purified and labeled by Paul Lauterjung. Using ensemble-based Time Correlated Single Photon Counting (TCSPC) methods we resolved two distances for every FRET pair, for M_1 and M_2. These distances were used in combination with the existing data to screen an initial ensemble of structures for M_1 and M_2 generated from a combination of Molecular Dynamic (MD)-simulations, rigid body docking and clustering methods [84]. To do so, the previous resolved FRET pairs $N_{FRET,previous}$ = 12 were taken and combined with the additional $N_{FRET,new}$ = 11 and the same ensemble of distances were screened jointly with N_{FRET} = 23. As an outcome, the structural ensemble width described using an average RMSD narrows down from 11.2 Å to 6.3 Å for the M_1 state of the non-farnesylated hGBP1, and from 14.5 Å to 8.2 Å for the M_2 state (see Figure 27 a, c). For specific samples labeled at the α13-helix of the hGBP1 we found an additional, static species in the data. Hence, we used the third distance to screen the previously derived M_1 ensemble using a χ^2-threshold. The result is a sub-ensemble of structures where the α13-helix is significantly closer in respect to the α12-helix (see Figure 27 b).

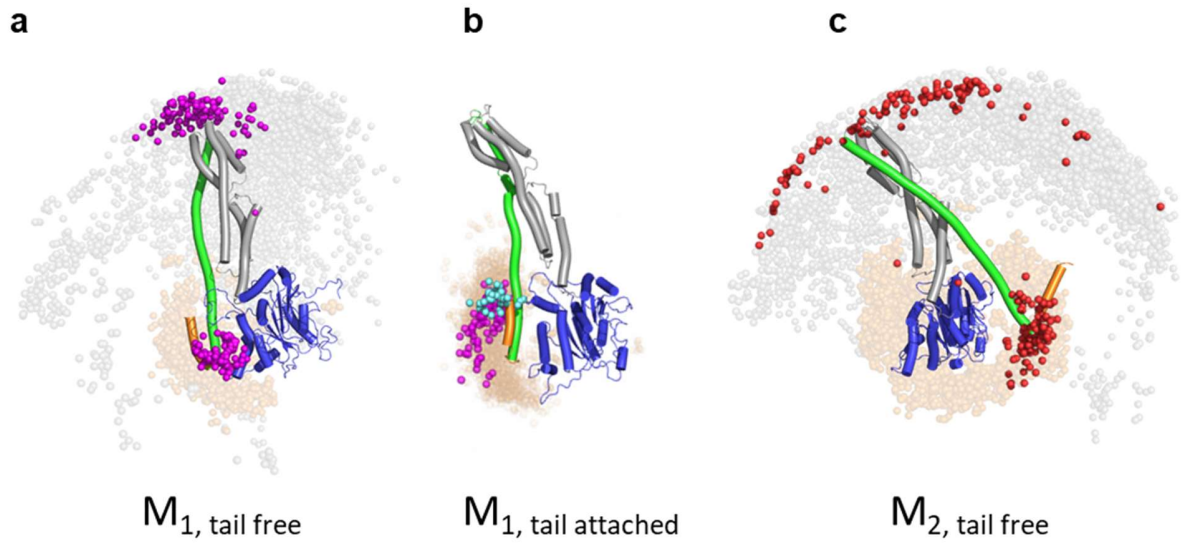


Figure 27: Structural refinement of hGBP1. a Structural ensemble of hGBP1 monomeric state 1 using a Meta-Analysis of EPR, SAXS and FRET including N_{FRET} = 23 (left) distance measurements. Structures are aligned to each other and magenta spheres indicate structures with a p value = 0.68 discriminating 95% of all structures. Magenta spheres represent the amino acids T481 and F565 of the $M_{1,tail\ free}$ ensemble. Full ensemble with rejected structures is shown in transparent gray and orange colors. Best structure is shown as cartoon. **b** Sub-ensemble of $M_{1,tail\ free}$ with spheres indicating the last amino acid of the α13-helix (M582) for the sub-ensemble (cyan) and the $M_{1,tail\ free}$ ensemble (magenta). **c** Ensemble of the $M_{2,tail\ free}$ state, with red spheres representing the best structures.

Complex dynamics between α12- and α13 helix in hGBP1$_{non\text{-}farn}$

Using two distinct samples (E533C/M583C, E521C/K582C) it was possible to monitor the movement of the α12 (labeled at E533C and E521C) in respect to the α13 helix (labeled at E583C and E582C). As a result, we found that in the non-farnesylated hGBP1 the α13 helix shows a

diffusive behavior, meaning it can open the angle between the α12/13-helices and diffuses on a time scale of a few µs (example measurement and explaining sketch shown in Figure 28). This has not been resolved previously because the newly used FRET pairs show a very high FRET efficiency contrast to this specific movement. Different methods were used to investigate this movement, namely *Efficiency-tau* plots using FRET lines, Photon Distribution Analysis (PDA), filtered Fluorescence Correlation Spectroscopy (fFCS) and sub ensemble TCSPC. Combining these methods, it was found that the molecule follows a quasi linear model with an additional static state. More specifically, we found that a small fraction of the $hGBP1_{non\text{-}farn}$ is in the static tail-attached $M_{1,ta}$ state, where the α13 is attached to the protein, for which electrostatic interaction might be the reason. The majority of the protein is in the dynamic exchange between the tail free $M_{1,tf}$ and $M_{2,tf}$ states, for which each of the states the diffusing behavior of the α13-helix comes on top. Thus, molecules also show a multi-state kinetic with three FRET states involved. This can be in a more accessible way analyzed using recently developed theory described in [33] which is based on the representation of the FRET lines using the difference between the normalized first and second moment Γ which is defined as

$$\Gamma = \frac{\langle \tau_{D(A)} \rangle_x}{\tau_{D(0)}} - \frac{\langle \tau^2_{D(A)} \rangle_x}{\tau^2_{D(0)}} \tag{4.3-1}$$

with the fluorescence lifetimes τ of the donor and acceptor (D(A)) and only the donor (D(O)) and the species weighted value $\langle\,\rangle_x$. This can be also represented using experimental observables resulting to

$$\Gamma = (1 - E)E_\tau, \tag{4.3-2}$$

with $E_\tau = 1 - \frac{\langle \tau_{D(A)} \rangle_F}{\tau_{D(0)}}$, where $\langle\,\rangle_x$ denotes the fluorescence weighted lifetime. Using this representation, the static FRET line transforms to a parabola defined by

$$\Gamma_{static} = (1 - E)E, \tag{4.3-3}$$

and, more interestingly, the dynamic FRET line is a straight line connecting the two limiting states (1) and (2) via

$$\Gamma_{dynamic} = \left(1 - E^{(1)} - E^{(2)}\right)E + E^{(1)}E^{(2)}. \tag{4.3-4}$$

Using this I could identify the contributing two state dynamics, but also found, as expected from our model, a region with a three-state dynamic between the dynamic lines $\Gamma_{dynamic}$. I denote that this is the first time one could use this theory to apply it to data describing a biomolecule. It can be summarized that additionally to the α12-helix dynamic a dynamical behavior of the α13-helix is found resulting in the $M_{1,tf}$ and $M_{2,tf}$ states.

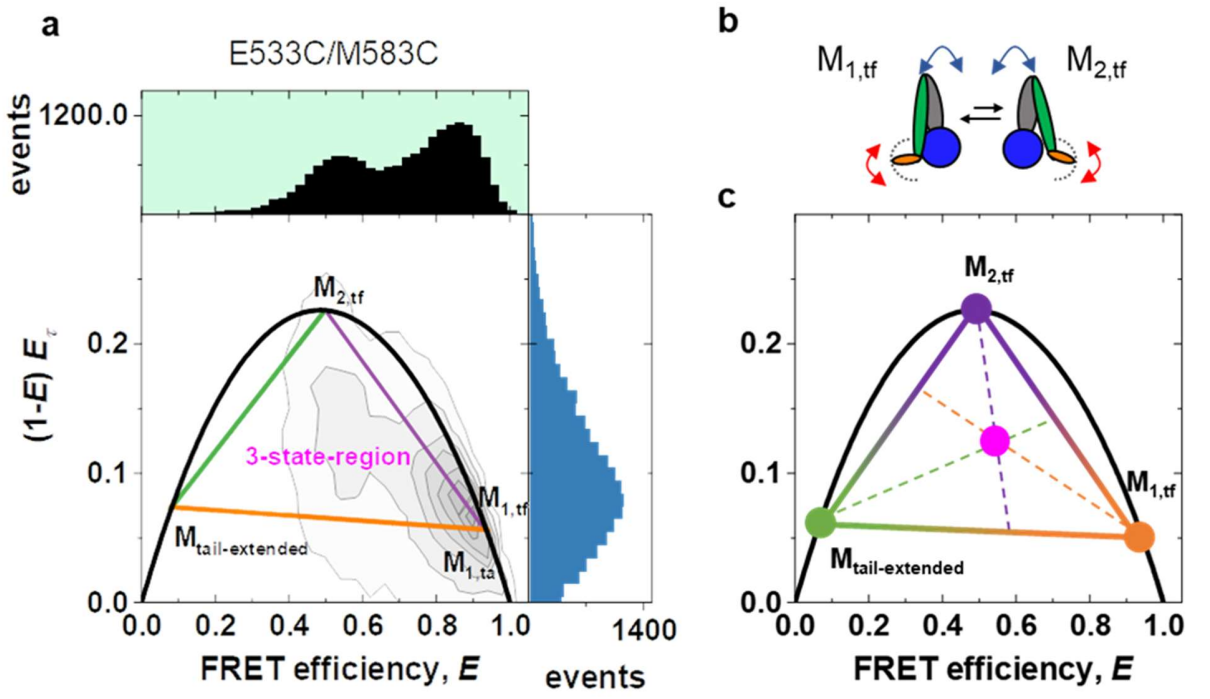


Figure 28: Key measurement of hGBP1 monitoring the α12- and α13-helix. a 2D sm-MFD histogram of E533C/M583C sample labeled with Alexa488 as donor dye and Alexa647 as acceptor dye. Top of histogram shows the distribution of FRET efficiency values. Right of histogram shows the difference between the normalized first and second moments of the lifetime distribution as further described in the text and in [33]. In the 2D histogram only bursts are shown which are originating from double labeled molecules, filtering was done by using a stoichiometry, ALEX-2CDE and TGX cut. Black line represents static FRET line as a parabola as further described in [33], dynamic FRET lines are shown as straight lines in purple for $M_{1,tf} \rightleftharpoons M_{2,tf}$, in green for $M_{2,tf} \rightleftharpoons M_{tail\text{-}extended}$ and orange for $M_{1,tf} \rightleftharpoons M_{tail\text{-}extended}$. **b** Sketch of the non-farnesylated hGBP1 in its $M_{1,tf}$ (left) and $M_{2,tf}$ state (right) showing complex dynamics with a rolling of the α-12 helix (blue arrow) and a diffusional behavior of the α-13 helix (red arrow). **c** Sketch of a three-state system and the resulting FRET lines. It is based on a recent study about FRET lines [33]. Orange sphere represents the $M_{1,tf}$ state, purple sphere the $M_{2,tf}$ state, green the tail extended state and magenta the multiple state area, where the obtained signal is a mixture of the three resolved states. Color of the lines indicate the current state of the molecule.

Farnesylated hGBP1

When measuring the farnesylated hGBP1, we found a drastic change in the dynamical behavior. While, as described before the non-farnesylated hGBP1 showed complex dynamics in its monomeric state, the $hGBP1_{farn}$ showed only a static state, which is defined as $M_{1,locked}$. Using a FRET label network connecting all domains of the $hGBP1_{farn}$, only the $M_{1,locked}$ state appeared in the data, which had for the E533C/M583C sample a slightly higher FRET efficiency value indicating a closer position of the α13-helix in respect to the α12-helix. These findings showed that the farnesyl moiety causes a dramatic change in terms of dynamics and structures. Since the hGBP1 performs a whole pathway from the monomeric to the oligomeric state in order to fulfil its biological functions, further experiments aimed to discover this pathway for the $hGBP1_{farn}$. Thus, a nucleotide is added to the solution and curiously, no changes could be observed (see Figure 29, left part). The $hGBP1_{farn}$ showed no change and was still in its sole resolved state. After addition

of unlabeled wild type $hGBP1_{farn}$ in the low µM range, the dynamical behavior identical to the one which were found in case of the $hGBP1_{non\text{-}farn}$ reoccurs. We interpret this result such, that the $hGBP1_{farn}$ dimerizes via LG:LG contact, and activates its dynamic as a search process to form the more stable, bridged dimer with an additional contact of the α13-helix (see Figure 29, middle part). After an additional waiting time of approximately 20 minutes at the used molecule concentrations, large structures are found, indicating that the molecule oligomerized. It is assumed, that before oligomerizing it forms a stretched dimer (see Figure 29, right part), since the data also showed a stretched state while it was still in its dimeric state. A kinetic analysis showed that this state is only slowly populated, while for the $hGBP1_{non\text{-}farn}$ dimerization is fast. The closed dimer (see Figure 29, middle part) could not be found, suggesting that this state is a rather fast decaying transition state. All used samples were measured in its oligomeric state and a simple approach based on interpreting FRET measured distances as a 3-dimensional vector was used to compare the oligomeric state of the $hGBP1_{farn}$ (for details see manuscript) to literature values [31, 87]. Here, other techniques were used to estimate the full stretched size of a single $hGBP1_{farn}$ in its oligomeric structure. As a result, I found good agreement and interpret this as that the α12/13-helices are fully elongated within the oligomeric $hGBP1_{farn}$.

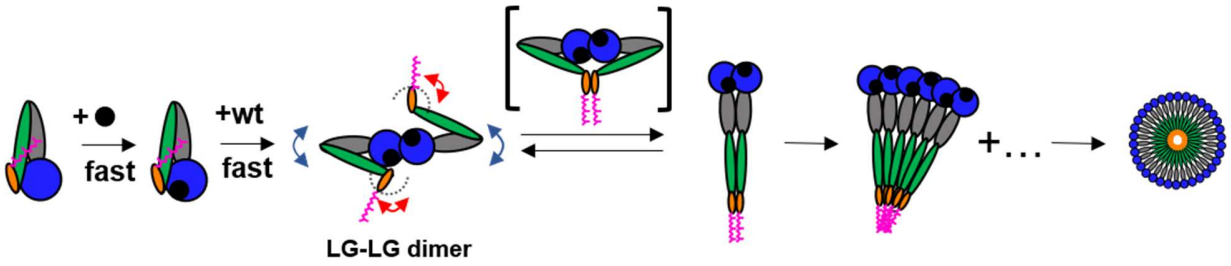


Figure 29: Oligomerization pathway of the farnesylated hGBP1. Sketch of the oligomerization pathway of the $hGBP1_{farn}$ under addition of a nucleotide (black sphere) and additional farnesylated wild type (wt). Blue arrows indicate the dynamic movement of the α12-helix, red arrows of the α13-helix. After binding of a nucleotide, the molecule does not change. Only by adding further monomers (here wild type is used), the dynamics resolved previously reoccurs. After passing a fast transition state the molecule slowly populates the state of an extended dimer. As a next step, a concentration dependent formation of the oligomeric structure is found.

4.3.4 Conclusion

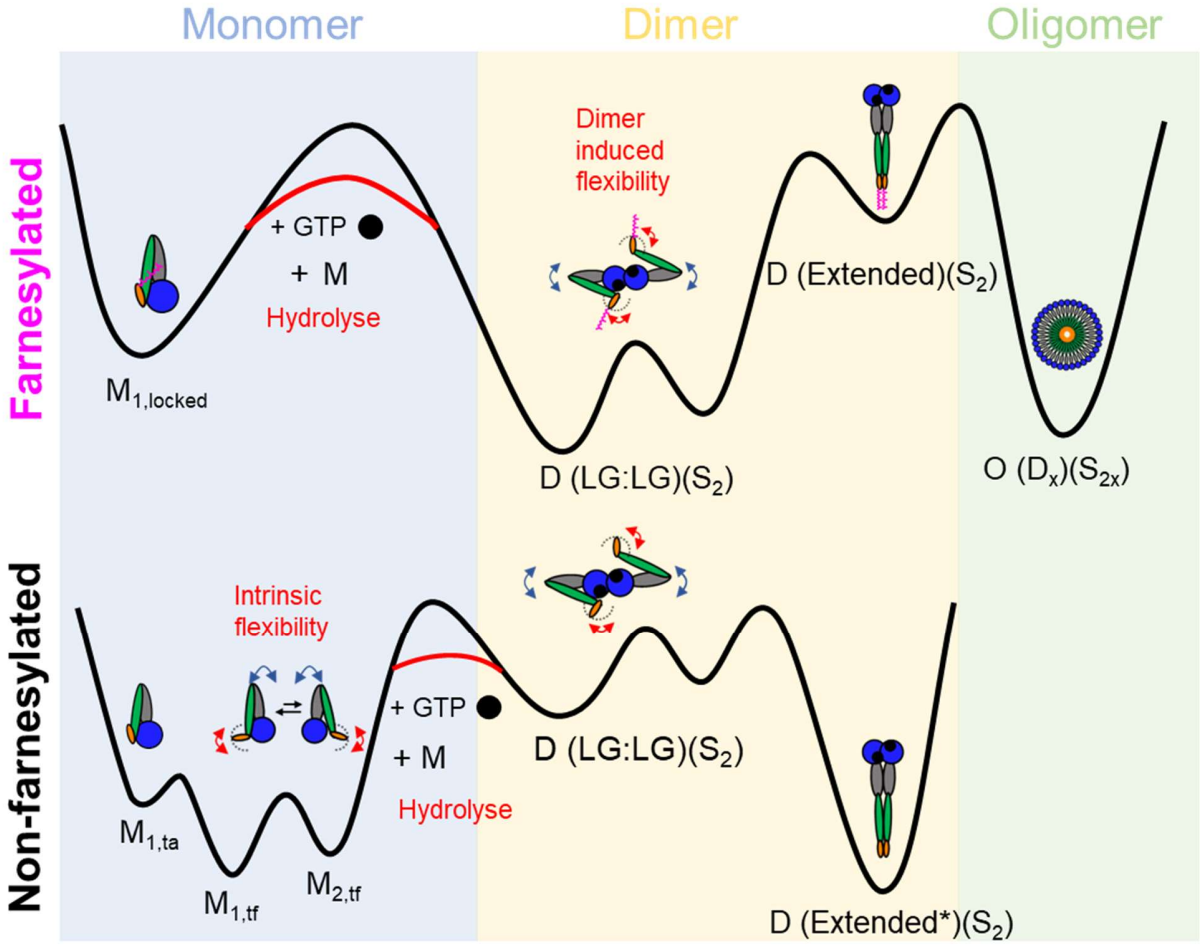


Figure 30: Comparison of energy landscapes of the non- and farnesylated hGBP1. Blue area indicates the monomeric (M) state of the non- and farnesylated hGBP1, yellow the dimeric state and green the oligomeric state. In the monomeric (M) state, only the non-farnesylated hGBP1 shows a dynamical behavior (intrinsic flexibility), which is triggered at a later stage of the farnesylated hGBP1, which is after dimerization (dimer induced flexibility). The extended dimer (D) of the farnesylated hGBP1 is slower populated compared to the non-farnesylated hGBP1, hence the depth of the energy potential is much less. S_x indicates the number of bound substrates, which is GDP-AlF_x. No detailed information is given to determine if their structures are identical, hence the * in the non-farnesylated case. The final step, the oligomerization can only be performed by the farnesylated hGBP1.

In this study it was found that the $hGBP1_{non\text{-}farn}$ possesses complex dynamics which mainly involve a movement of the α12-helix resulting in two monomeric states $M_{1,tf}$ and $M_{2,tf}$ and with a diffusing α13-helix. Due to the diffusing α13-helix the $hGBP1_{non\text{-}farn}$ can already reach positions for its tail that it needs at a later stage, which is the extended dimer and for the farnesylated hGBP1 the oligomerization. On top, we found a minor populated static fraction of the $hGBP1_{non\text{-}farn}$, where the α13-helix is attached to the α12-helix ($M_{1,ta}$). For the $hGBP1_{farn}$ the molecule does not show any dynamical behavior at all until a nucleotide and additional wild type is added into the solution. From that point on, the same dynamics shown by $hGBP1_{non\text{-}farn}$ monomers were resolved while monitoring the molecule in its dimeric state. This important finding is answering the question of

the driving force of the hGBP1, which is a state dependency and therefore concentration dependent (see Figure 30). As a result of dimerization, the farnesyl moiety is released and dynamics are activated (dimer induced dynamics). Its purpose is to form a more stable state, which is the bridged dimer. After that, the dimer slowly populates the extended dimer state via passing the bridged dimer as an intermediate state. This finding is remarkable since this shows another crucial difference to the non-farnesylated hGBP1, as the extended dimer is formed fast as previously seen by T. Peulen [unpublished data]. As a final step, only the $hGBP1_{farn}$ shows a concentration dependent oligomerization.

Using this information, we found that the hGBP1 possesses controls that are relevant at later structural stages of it. For the $hGBP1_{non\text{-}farn}$ we found dynamics in the monomeric state that are relevant after dimerization of the $hGBP1_{farn}$ releasing its farn anchor and unlocking its only static state $M_{1,locked}$. On top we found that the $hGBP1_{non\text{-}farn}$ can already stretch its tail to a full extended state, relevant for the elongated dimer of the $hGBP1_{non\text{-}farn}$ and $hGBP1_{farn}$ and the oligomeric structure of the $hGBP1_{farn}$. We can conclude that the hGBP1 is a molecular machine that possesses relevant controls already in its monomeric state defining how it performs its biological function.

4.4 Resolving the mechanism behind pore formation of the Tc toxin via an analysis of states, rates and transition times

Roderer, D. [1,a,#], Kucher, S. [2,b,#], Folz, J.[3#], Njenga Ng'ang'a, P. [1], Kühnemuth, R.[3], Assafa, T.E.[2], Bordignon, E. [2,b,*], Seidel, C.A.M.[3*], Raunser, S. [1,*]

Author list preliminary.

[1] *Department of Structural Biochemistry, Max Planck Institute of Molecular Physiology, Dortmund, Germany*

[2] *Faculty of Chemistry and Biochemistry, Ruhr University Bochum, Bochum, Germany*

[3] *Chair of Molecular Physical Chemistry, Heinrich-Heine University Düsseldorf, Germany*

[a] *Current address: Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP), Berlin, Germany*

[b] *Current address: Department of Physical Chemistry, University of Geneva, Switzerland*

[#] *These authors contributed equally*

[*] *Corresponding author*

The following chapter is based on a collaboration managed by S. Raunser from the Max Planck Institute of Molecular Physiology in Dortmund. Here, the contribution of the fluorescence spectroscopy experiments is summarized with schematic models developed by D. Roderer and S. Raunser. Since the manuscript is in writing phase, further data from cryo electron microscopy and EPR measurements is excluded.

4.4.1 Background: Molecule of interest

Heterotrimeric toxin complexes (Tc) are virulence factors of many bacteria including insect and human pathogens [88, 89]. They perforate the membrane of the host and inject a toxic enzyme to the cytoplasm of the host cell. The reaction of the enzyme leads to a concrete actin clustering in the cell and finally to its death [90]. Tc toxins were originally discovered in the insect pathogen *Photorhabdus luminescens* [91], but appear to be represented in enterobacteria as some insect pathogens and facultative human pathogens e.g. Salmonella enterica and Yersinia pseudotuberculosis as well [92-96]. Tc toxins of insect pathogens are candidates for biopesticides and currently the focus of crop protection research [93, 97], while Tc toxins of human pathogens are medically relevant [98, 99]. The Tc toxin consists of three components, which is TcA, TcB and TcC. The TcA is an around 1.4 MDa pentamer with five equal subunits (see Figure 31 and Supplementary Figure 10). It has an outer shell with an inner pore. The TcA forms a complex with TcB-TcC, which together form a hollow cocoon with the encapsulated toxic enzyme, coded by the Hypervariable Region (HVR).

4.4.2 Aim of the study

The structure of the Tc toxin was studied in recent years. In 2013/14 a syringe-like injection mechanism was found [100], with start- and end-structures resolved in molecular detail [101]. As a next step, the membrane insertion including its involved structures was investigated [102, 103].

Furthermore, the glycan-dependent cell adhesion was resolved [104]. As a next aim of investigation, the conformational changes the molecule performs should be investigated (see Figure 31a). In case of a pH shift from 7.5 to 11.2 the molecule performs a dramatic conformational change from a PrePore state to a Pore state, which has not been resolved yet. This conformational change includes that the molecule opens its shell and injects its pore. Since this transition involves large conformational changes, this study addresses the following points to resolve the mechanism behind this transition:

- kinetic analysis of the shell opening and pore injection processes, i.e., resolving rates of each process and develop a kinetic model (1)
- investigation of potentially stable and transient intermediate states (2)
- influence of mutations and other perturbances like biotinylation (3).

Using this information, we want to resolve the energy landscape of the Tc toxin (see Figure 31b) in order to understand its mechanism. To accomplish this, we designed different types of experiments using single-molecule fluorescence techniques. This involves confocal fluorescence spectroscopy as well as microscopy methods. Since the TcA is a pentamer with five structurally equal subunits, precise labeling is very challenging. Hence, we designed different assays and samples using homo- and heteroFRET to address the raised questions. The samples were prepared by D. Roderer, P. Njenga Ng'ang'a and S. Raunser.

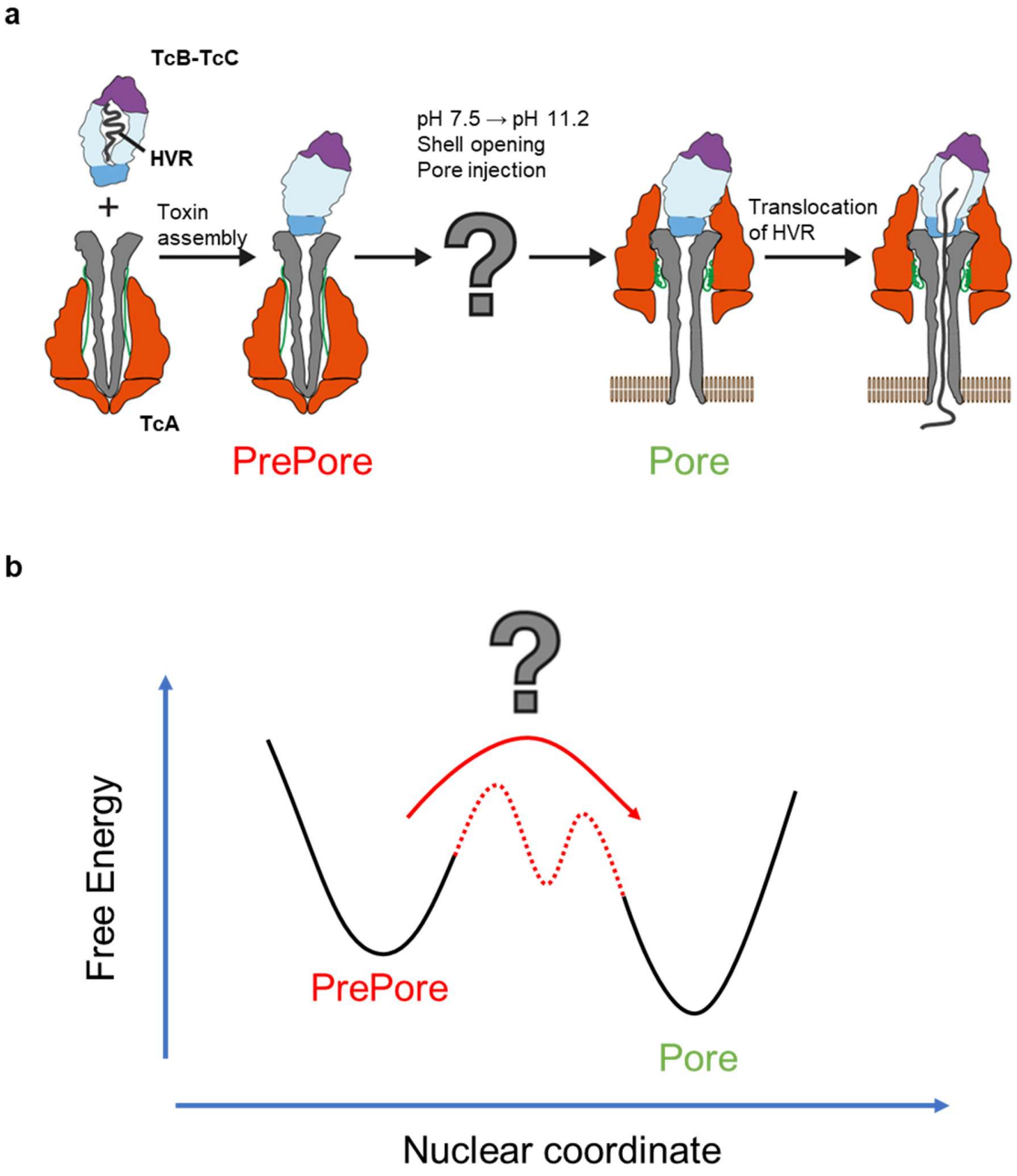


Figure 31: Schematic representation of the syringe like mechanism of the Tc toxin. a Scheme created by D. Roderer and S. Raunser. The shell of the Tc toxin is shown in orange, with the pore channel in grey. TcA consists of the shell and the channel. TcB-TcC complex is shown in blue, light blue and purple. The Hypervariable Region (HVR) is located inside of TcB-TcC and encapsulates the toxin component. After shell opening and pore injection to a membrane in a syringe like mechanism the toxin can be released into a cell. **b** Energy landscape representation of the transition from the PrePore state to the Pore state. It is unknown if the molecule passes intermediate states or has a direct transition.

4.4.3 Results

For all following results, method descriptions can be found in the supplementary information under chapter 6.2.

Kinetic analysis of Tc Toxin using smFRET measurements

We designed the initial experiments to study the kinetics of the pore formation of the Tc toxin. The molecule performs a major conformational transition upon a shift of the pH from neutral to alkaline (pH 11.2) in the solvent. During this transition, the Tc toxin opens its shell and consecutively a pore is injected. To probe for shell opening, the TcA of the Tc toxin was labeled at positions 1193 with BodipyFL-iodacetamide (see Figure 32a, top). Since the Tc toxin is a homopentamer, it consists out of five equal subunits, which are each labeled at position 1193. Due to this, heteroFRET derived distances are very challenging to measure, because precise labeling is not possible. Hence, we use homoFRET which affects only the polarization of the emitted photons which is changed in case of homoFRET. It is expected that after the shell is opened, an increase in the residual anisotropy can be observed since the dyes will move away from each other which will result in a reduction of the homoFRET efficiency. This is because if homoFRET is present the emitted photons are depolarized due to the transfer from one dye to another dye. For the pore injection, fluorescent dyes BFL and At647N were introduced at the positions 2365 and 2384, respectively. Here, the dyes are expected to be within the inner part of the shell after the conformational change (see Figure 32a, bottom). This will result in an increase of the residual anisotropy due to the restricted movement of the dyes in the pore state. As a method, we used single-molecule counting and analyzed the data using Photon Distribution Analysis (PDA). More information about the method can be found in the supplementary information chapter 6. Underlying data can be found under Supplementary Figure 1 and Supplementary Table 2.

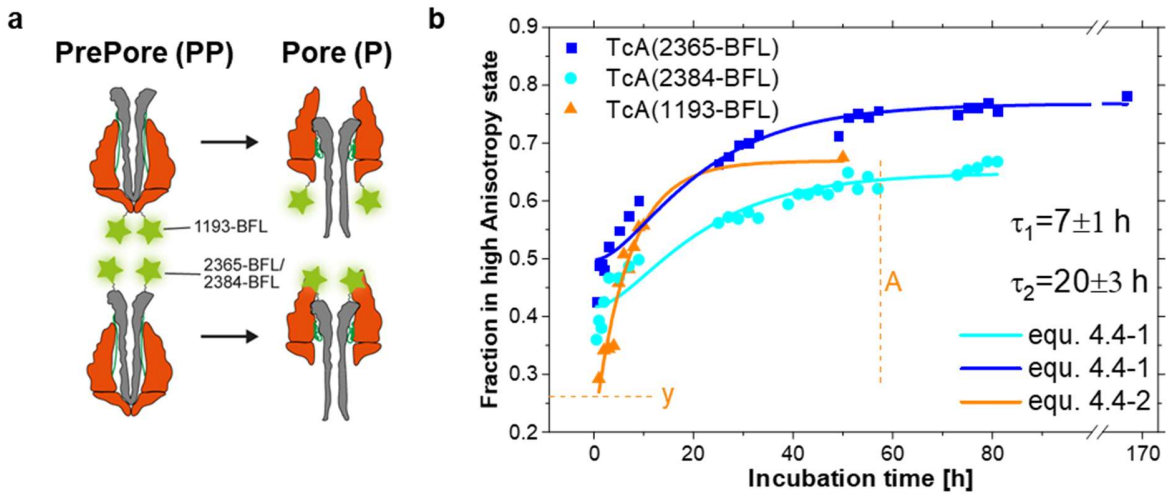


Figure 32: Kinetic analysis of the conformational change from PrePore to Pore state of the TcA. a Schematic representation of the PrePore (PP) and Pore (P) state with the involved label (BodipyFL) at positions 1193 to monitor the shell destabilization and positions 2365 and 2384 to probe the pore injection. **b** Fraction of the high anisotropy state obtained in PDA for the shell destabilizing sample (orange) and the pore injection samples (blue, cyan). Each data point represents a one-hour time window. Fractions are fitted using a consecutive kinetic model leading to two relaxation times shown as inset (see equ.(4.4-1) and (4.4-2)). *A* represents the total amplitude of the anisotropy change and *y* the offset. For 1193-BFL the high anisotropy state (r = 0.28) is caused by environment change and reduction of homoFRET, indicating a change of the shell. In the case of 2365-BFL and 2384-BFL high anisotropy state (r = 0.29, and r = 0.30) is caused by environment change from the surrounding opened shell/injected pore therefore interpreted as Pore state.

To fit the obtained time dependent fractions of the high anisotropy state, which is due to a distance increase of the dyes, we used a consecutive kinetic model. The reason is that a clear difference in the kinetic rates was visible, since the samples labeled to monitor the shell opening showed higher rates than the samples monitoring the pore injection. Hence, we concluded that an intermediate state needs to be populated first before the pore can finally be injected. One could also explain it the other way around: If the pore were to be injected immediately after the shell opens, no difference in the kinetic rates would be seen since the transition path time of the molecule can be neglected. This potential hypothesis is falsified since the measured rates differ significantly. Therefore, data was fitted using following consecutive model:

$$[Stable\ Intermediate\ 2] = A\left(1 - e^{-\frac{t}{\tau_1}}\right) + y \tag{4.4-1}$$

$$[Pore] = A\left(1 - \frac{\frac{1}{\tau_2}e^{-\frac{t}{\tau_1}} - \frac{1}{\tau_1}e^{-\frac{t}{\tau_2}}}{\frac{1}{\tau_2} - \frac{1}{\tau_1}}\right) + y \tag{4.4-2}$$

where τ_1 describes the relaxation time of the PrePore to destabilize its shell and τ_2 the relaxation time of the shell destabilized PrePore to transform into a Pore by injection of the channel. *A* is the total amplitude of the anisotropy change and *y* the offset. For the sake of generality, the intermediate is already called Stable Intermediate 2 (SI-2), for which reason will be introduced in the later sections. It can be seen that the SI2 state is populated with a relaxation time of τ_1 = 7$\pm$1 h and the injected pore with τ_2 = 20$\pm$3 h. These results were confirmed using other types of experiments as well, which are electron paramagnetic resonance (EPR) and electron microscopy (EM), but cannot be discussed here due to the current unpublished state of the manuscript. We can conclude that after the pH change, the shell of the toxin destabilizes and opens faster as the pore is injected, hence a stable intermediate of an open or partially open shell is populated.

Time resolved anisotropy decays resolving an intermediate state

To further investigate the resolved intermediate state (SI-2), a new mutant was developed in the group of S. Raunser. By mutating the lysine at the position 1197 to a tryptophan (variant K1179W), the Tc toxin loses its ability to inject the channel and will possibly only populate the found intermediate state, which could then be studied in greater detail. For further testing of the change of distances, based on homoFRET samples were labeled with a high and a low amount of dyes aiming for a fully labeled pentamer and a pentamer with only one labeled subunit, respectively. In the case of low labeled samples, no change of the homoFRET efficiency is expected since it is based on the interaction of multiple dyes.

As seen in Supplementary Figure 4 and Supplementary Figure 5, the sample labeled at the position 2365 to monitor the pore injection did not exhibit any signal change as a function of time at all. This confirmed the expectation that the mutant K1197W is not able to reach the final Pore state. As it can be seen in Supplementary Figure 2 and Supplementary Figure 3, the mutant showed a slight change in the anisotropy for the high labeled samples and, on top, a change in the fluorescence lifetime with a relaxation time of $\tau_0 \approx 10$ min. Since the single labeled sample did not show a change in the anisotropy, the observed change for the high labeled sample is originating from a distance change, but compared to the one measured using the wild type significantly weaker, as it will be discussed next. The lifetime change occurred due to dynamic

quenching effects (see chapter 3.4) resulting from the tryptophan, which seem to be less intense after the conformational change.

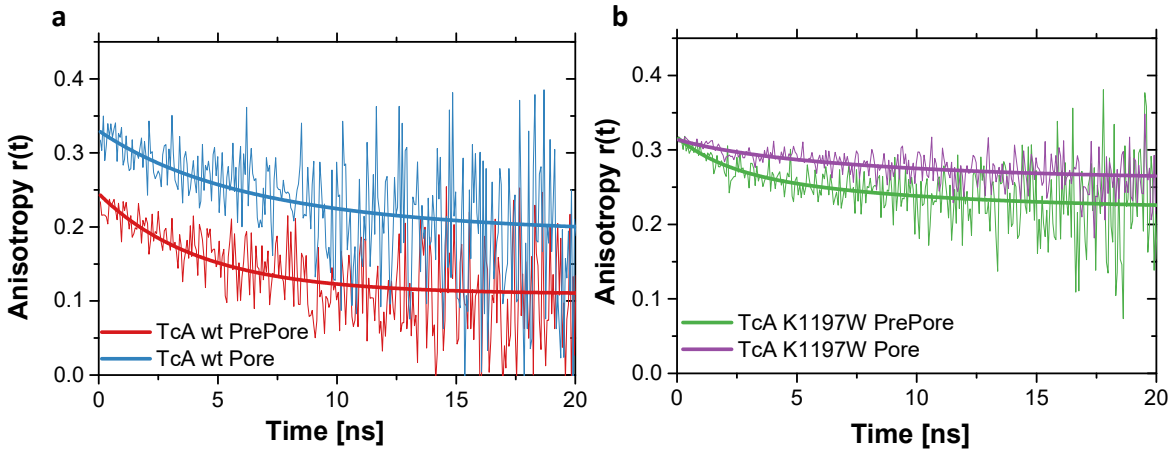


Figure 33: Time resolved anisotropy decays of Pore and PrePore states of the TcA. a Blue is showing the time resolved anisotropy decay fitted using equation (4.4.3-3) and (4.4.3-4), obtained via smMFD measurements for the wild type Pore state at the end of the pH 11 measurement, and in red for the wild type PrePore state at the beginning of the pH 11 measurement. **b** The same curves are created for the K1197W mutant, showing the time resolved anisotropy decay of the Pore (purple) and the PrePore (green).

To quantify the distance change further and characterize the intermediate state the mutant K1197W populates, time resolved anisotropy decays, *r(t)*, were extracted and fitted using theory described in previous literature [47] (see Figure 33). As a result, fitting led to two depolarizing constants with a lifetime of fast depolarizing (2-4 ns) and slow depolarizing (9 ns), with their amplitudes A_{fast} and A_{slow}, respectively. Since *r(t)* should start at the fundamental anisotropy r_0= 0.38, another amplitude $A_{very\ fast}$ was added to the fast depolarizing amplitude which is the difference between the sum of the resolved amplitudes $A_{depolarization}$ and the remaining difference of r_0 and the offset y_0:

$$A_{depolarization} = A_{very\ fast} + A_{fast} + A_{slow} = r_0 - y_0 \qquad (4.4.3\text{-}3)$$

$$A_{very\ fast} = r_0 - y_0 - A_{fast} - A_{slow} \qquad (4.4.3\text{-}4)$$

As it can be seen in comparison of Figure 33 a and b, the difference of depolarization between the PrePore and the Pore state using the wild type sample is quantitatively larger than the one using the K1197W sample, indicating a higher distance change in the wild type sample. To add more reliability to these findings, we predicted the change of the mentioned amplitude based on the Degree of Labeling (DoL) of the used samples and the structural models which are resolved by D. Roderer. To estimate the DoL we used FCS to analyze the brightness and compared it to single labeled samples and free dye measurements, where Rhodamin110 is used. It is seen that labeling of the K1197W sample did not work very well and only an average dye number of N_{dye}= 1 was measured for the sample with the high amount of dyes (whereas we measured for the wt sample an average dye number of N_{dye}= 2.5) resulting in a reduced number of molecules that can undergo homo FRET. However, an average dye number of 1 is still resulting in a high number of

molecules with having more than two dyes labeled based on Poissonian statistics (to give numbers: for an average number of 1 dye per molecule, 40 % of the molecules have more than 1 dye attached). As a next step, FRET rates were calculated for all possible labeling schemes and probabilities based on the labeling probability distribution with the assumption that in case of multi labeling the FRET rates simply add up (sketch can be found under Supplementary Figure 6, distribution of dyes under Supplementary Figure 7 and more detailed results in Supplementary Table 3). Comparing the expected FRET active molecules (number of dyes greater or equal than 2) from the prediction to the experiment showed a good agreement, indicating that our assumptions worked (see Figure 34a). Interestingly, for the change in the fraction of the depolarizing through homo FRET (see Figure 34b) we found an agreement between experiment and prediction for the wild type sample, whereas we could not detect a high change in the experiment of the K1197W sample, mismatching the predictions. We concluded that the distance change in the K1197W sample is only minor and we do not see an open shell for this molecule. This was confirmed by D. Roderer, who found out that this new intermediate is formed by a movement of the receptor-binding domain B (RBD-B) into the accessible volume of the dye, resulting in a change of anisotropy. This was done by acquiring a cryo-EM map through 3D classification showing higher densities at the suspected position for the RBD-B. The model was fitted using a rigid body fit. Summing up, using the K1197W sample, another stable intermediate 1 was found, which is prior and different to the stable intermediate 2.

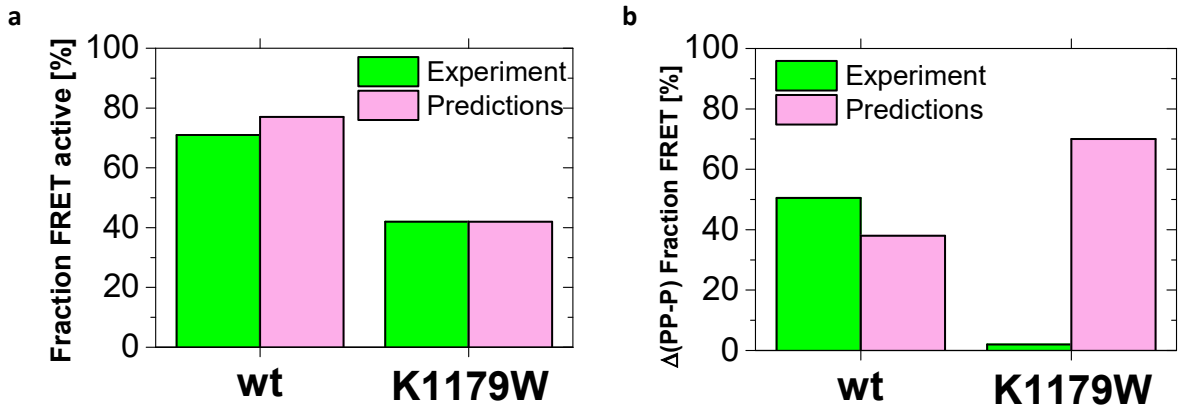


Figure 34: Experiment to predictions comparison. a Comparison for the fraction of homoFRET active molecules resulting from the experiment (green) to the predictions (magenta). Left 2 columns show the values for the TcA wild type sample, right 2 columns for the TcA K1197W sample. **b** Comparison for the experimentally resolved and predicted change of fraction due to the conformational change from PrePore to the Pore state for the wild type (left) and K1197W sample (right).

pH change induces conformational flexibility sensed by polarization resolved FCS (pFCS)

As the K1197W introduces a tryptophan in the proximity of the dye position 1193, sticking and quenching could be observed as discussed before. Even though these are effects that one usually wants to avoid (as discussed in chapter 3), the sticking of the dye can also carefully be used in an FCS analysis. In this approach, a correlation amplitude of a time t_c is calculated, which is called $G(t_c)$. The amplitude was fitted using the following equation:

$$G(t_c) = 1 + \frac{1}{N} \cdot \frac{1}{1 + \frac{t_c}{|t_d|}} \cdot \frac{1}{\sqrt{\left(1 + \frac{t_c}{(\frac{z_0}{\omega_0})^2 |t_d|}\right)}} \cdot \left(1 - |C| + |C| \cdot e^{-\left(\frac{t_c}{|t_C|}\right)} - |A| + |A| \cdot e^{-\left(\frac{t_c}{|t_A|}\right)} - |B| + |B| \cdot e^{-\left(\frac{t_c}{|t_B|}\right)}\right) \quad (4.4.3\text{-}5)$$

with *N* the average number of molecules in the observation volume, t_d the diffusion time, $\frac{z_0}{\omega_0}$ the gaussian shape factor, *C* the amplitude of the very fast detected dynamic, t_C the time constant of it, *A* the amplitude of the triplet and global rotation of the molecule, t_A its time constant and B the amplitude with time constant t_B of a third bunching term. Since the tryptophan is leading to a sticking of the used dye on the surface of the molecule, the dye is more sensitive to movements and flexibility of it. It can be seen that an anti-bunching term measured for the first 400 s of the measurement in pH 11 has a time constant of $t_{dynamic} \approx 100$ ns. After transition to the previously characterized SI1 state, this term is shifted towards much longer times of $t_{dynamic} \approx 500$ ns. Another time that could be observed is the global rotation of the molecule, which was estimated to around 4 µs, where simulation using the software HydroPro predicted 1 µs [105]. All fit results can be seen in Table 4. The fit was done globally over the parallel and perpendicular polarized autocorrelation and the cross correlation.

Table 4: Overview of fitted parameters for pFCS.

Sample	TcA(1193-BFL-K1179W) 0-400 s in pH 11	TcA(1193-BFL-K1179W) 4700-5200 s in pH 11	TcA(1193-BFL-K1179W) in pH 7
χ^2	3.3158	3.4052	4.0828
N	0.2591	0.5378	0.3988
t_d	4.3332	4.3332	4.3332
$\frac{z_0}{\omega_0}$	2.0567	2.0567	2.0567
C	0.5786	0.2272	0.3657
t_C	0.0001	0.0001	0.0001
A	0.1655	0.2137	0.1664
t_A	0.0041	0.0041	0.0041
B	0.0383	0.1093	0.1187
t_B	0.1683	0.1683	0.1683

Surprisingly, the correlation curve of the molecule in pH 7 shows the same, longer dynamic time. We conclude that the pH shift induces a flexibility of the TcA which is only present while the molecule is trying to find its stable intermediate 1, hence we call this state a Transient Intermediate (TI-1).

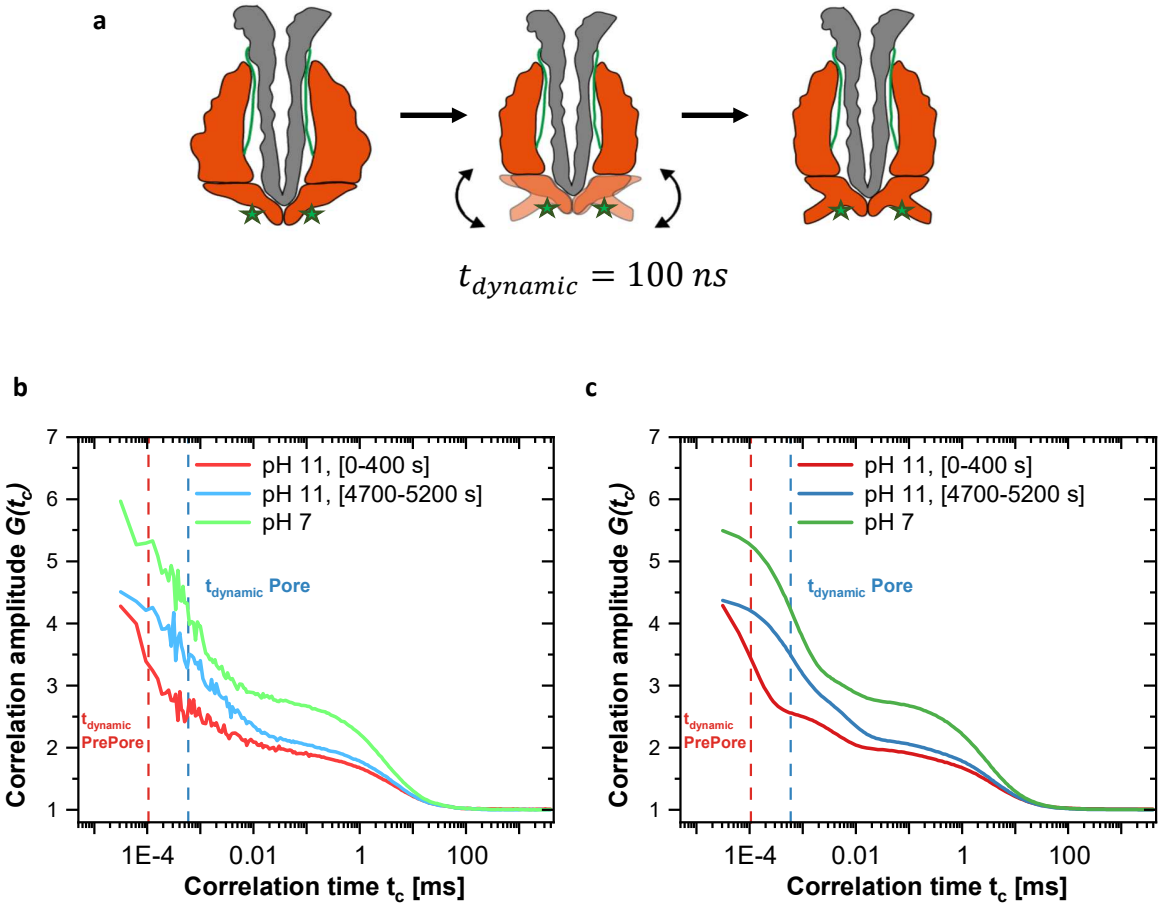


Figure 35: FCS of TcA K1197W labeled with BFL at positions 1193.a Scheme of the observed movement using FCS **b** Experimental data: y-axis shows the correlation amplitude $G(t_c)$ of the autocorrelation of the parallel polarized signals, x-axis the correlation time t_c. **c** Fitted correlation curves are calculated for the beginning of the pH 11 measurement (red), at the end of the pH 11 measurement (blue) and for the pH 7 measurement (green). Dynamic relaxation time $t_{dynamic}$ is indicated using dashed lines. In case of the PrePore state (red), $t_{dynamic} \approx 100$ ns, in case of the Pore state $t_{dynamic} \approx 500$ ns.

Transitions from PrePore to Pore resolved using TIRF

In the previous sections, mainly dwell times were studied which is the time during which the molecule stays in one conformation until it goes to another. As a side note: The average of the dwell times is the lifetime and in case of a one-directed dynamic the lifetime is the same as the relaxation time. To study the time the molecule actually needs to go from one state to the other, which will be called transition path time, another experimental approach is required. Since the previous experiments are based upon freely diffusional molecules, the transitions themselves occur in nearly all cases while the molecule is not in the confocal volume, due to relaxation times

of hours. Therefore, molecules need to be immobilized and “speeded up” which was achieved using Biotin (further details can be found in supplementary information in chapter 6). It could be seen that in case of the biotinylated Tc toxin the relaxation rates increased dramatically (see Supplementary Figure 8 and Supplementary Table 4). Using this effect, it became possible to use Total Internal Reflection Fluorescence (TIRF) microscopy, since the typical observing time of a molecule is between a few seconds and a minute until the dyes bleach.

To do so, two samples were designed to study the pore injection and shell destabilization, labeled and tested using smMFD experiments (smMFD test measurements see Supplementary Figure 9, design/distance predictions see Supplementary Table 5 and Supplementary Table 6, used 3D structures see Supplementary Figure 10). Here, homoFRET measurements were used to study the shell destabilization and heteroFRET measurements were used to study the pore injection (sketch see Figure 36). As a result, in case of the shell destabilization, a continuous increase of the polarization was observed after a relaxation time τ_1 (see Figure 36c, sketch see Figure 36e), ending after times of 30 ms up to 1.7 s. For the pore injection, an instantaneously, step wise increase of the FRET efficiency is monitored after a relaxation time τ_2 (see Figure 36d, sketch see Figure 36f) to the expected next neighbor FRET efficiency of approximately E = 0.25. All measured and more detailed traces can be found under Supplementary Figure 11 and Supplementary Figure 12. In total 765 FRET pairs were recorded with 48 of them showing a transition that could be tracked in the signal/trace. The rest remained unchanged due to the labeling scheme and potential bleaching of the dye. A histogram of all observed transitions (see Figure 36g) shows a broad distribution for probing the shell destabilization, whereas for the pore injection only transition times within a single bin could be observed. To analyze the measured relaxation times (see Figure 36h), the same consecutive model used before was applied, yielding to a similar ratio of τ_1/τ_2. We can conclude that while observing the transition from the closed shell to the open shell intermediates have to be passed leading to a broad distribution of the transition times. This confirms our findings from the previous experiments, which resolved a transient intermediate and two stable intermediates already. For the pore injection, no distribution of transition times was found, as expected. Here, the measured transition times to go from the PrePore state to the Pore state was below 60 ms.

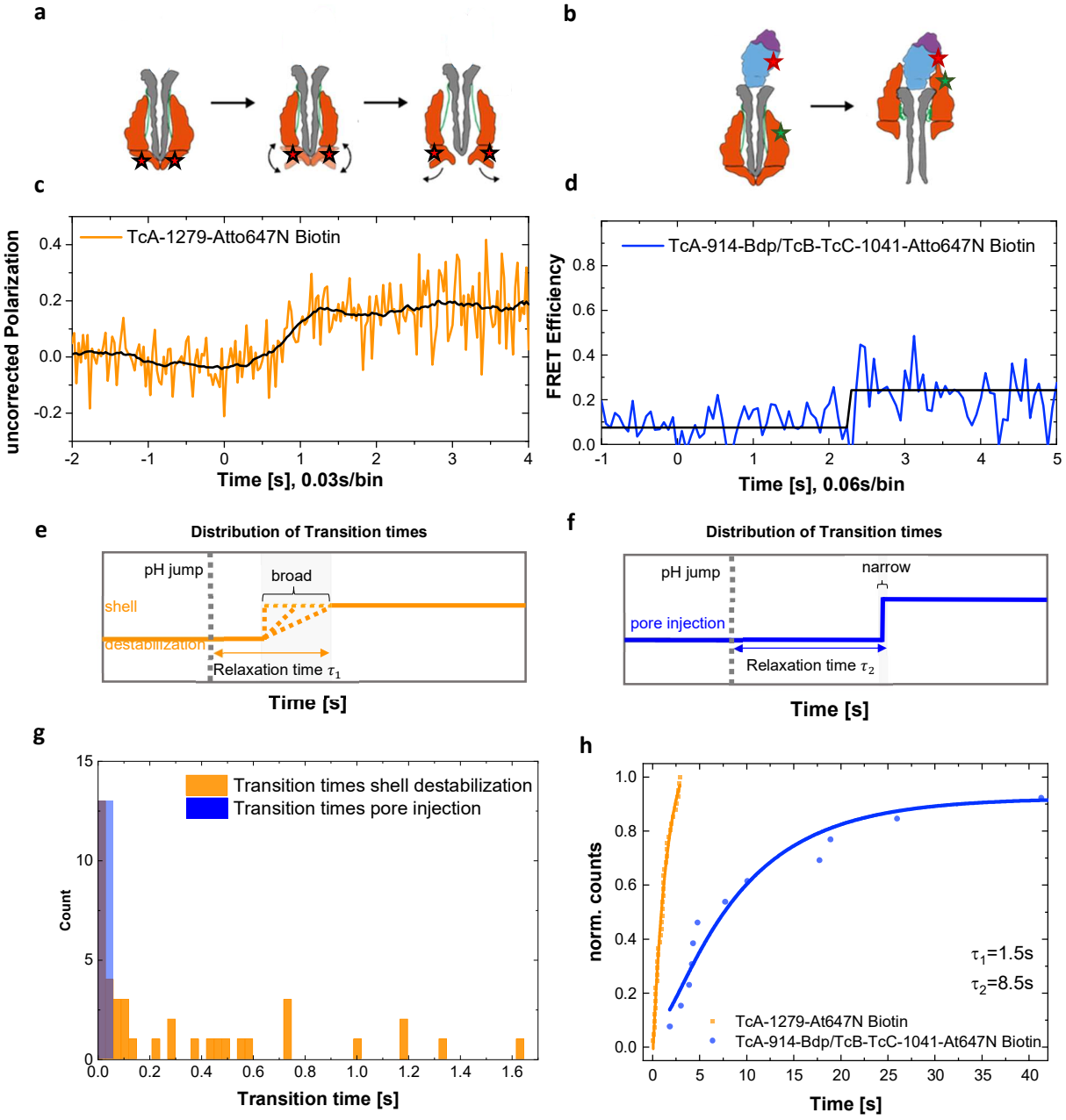


Figure 36: Summary of TIRF experiments. a, b Schematic representation of the sample probing the shell destabilization (**a**) and the pore injection (**b**), modified after a scheme created by D. Roderer and S. Raunser. **c** Example trace of monitoring the shell destabilization using a fully Atto647N labeled TcA-pentamer-1279. Higher transition times longer than the bin width (bin width = 29.44 ms) can be observed. The change in the polarization is caused by a difference in the homoFRET efficiency and reduction of mobility of the dyes due to moving binding domains. **d** Example trace of probing the pore injection using a donor dye at theTcA-pentamer-914 labeled with a dye to protomer ratio of 1:5 and acceptor dye labeled at TcB-TcC-1041. The predicted transfer efficiencies are $E_{\mathrm{PrePore}} = 0$ and $E_{\mathrm{Pore}} = 0.25$. The displayed trace is showing a fast transition of the signal from PrePore to Pore state within one bin (bin width = 58.88 ms). **e,f** Sketch of idealized signals indicating the obtained observables which are the relaxation times before the transition after pH change to the end of the transition and the transition time for the shell destabilization (**e**) and the pore injection (**f**). **g**

Histogram of the distribution of transition times for the shell destabilization (orange) and the pore injection (blue). **h** Normalized counts of relaxation times fitted with a consecutive model leading to $\boldsymbol{\tau_1}$ and $\boldsymbol{\tau_2}$.

4.4.4 Conclusion

Table 5: Overview of results describing state properties of the Tc toxin. Columns describe individual states involved in the PrePore to Pore transition with properties of the shell tip and the channel base (pore channel). Also listed are the used samples, methods and parameters. The shortage tr indicates the Time Resolved anisotropy, ss the steady-state anisotropy.

State:	PrePore	Transient intermediate 1	Stable intermediate 1	Stable intermediate 2	Pore
Shell tip (1193)	non-flexible	Flexible	small distance increase	large distance increase	large distance
Channel base (2363, 2384)	uncovered	uncovered	uncovered	uncovered	covered
Sample	TcA	TcA	TcA	TcA	TcA
Functional variants	K1197W and wt	K1197W	K1197W	wt	wt
Method	tr-anisotropy homoFRET, pFCS	pFCS	tr-anisotropy homoFRET	ss-anisotropy, tr-anisotropy homoFRET	ss-anisotropy
Parameter	interdye distance, anisotropy	anisotropy	interdye distance	interdye distance	mobility of dyes, anisotropy

Table 6: Overview of results describing the kinetics of the Tc toxin. Columns describe each rate involved in the PrePore to Pore transition k_0, k_1 and k_2 following the kinetic model shown in the first line. Also listed are the used sample, the probed process, the kinetic model used for fitting the PrePore or Pore decay/formation and used parameters. States are described as Transient Intermediate 1 (TI-1), Stable Intermediate 1 (SI-1) and 2 (SI-2), PrePore (PP) and Pore (P).

Kinetic model:	PP $\xrightarrow{k_0}$ TI-1 $\rightarrow$ SI-1 $\xrightarrow{k_1}$ SI-2 $\xrightarrow{k_2}$ P		
Result	$\frac{1}{k_0} = \tau_0 = 10$ min	$\frac{1}{k_1} = \tau_1 = 7$ h	$\frac{1}{k_2} = \tau_2 = 20$ h
Sample	TcA	TcA	TcA
Functional variants	K1197W	wt(1193)	wt(2365 and 2384)
Probing process	shell destabilization	shell destabilization and opening	pore injection
Kinetic model PP decay	mono-exp k_0 (1193)	mono-exp k_1 (1193)	-
Kinetic model P formation	-	-	consecutive fit with global k_1, k_2 (2365 or 2384)
Method	tr-anisotropy homoFRET, pFCS	ss-anisotropy, tr-anisotropy homoFRET	ss-anisotropy
Parameters	interdye distance, anisotropy	interdye distance, anisotropy	mobility of dyes, anisotropy

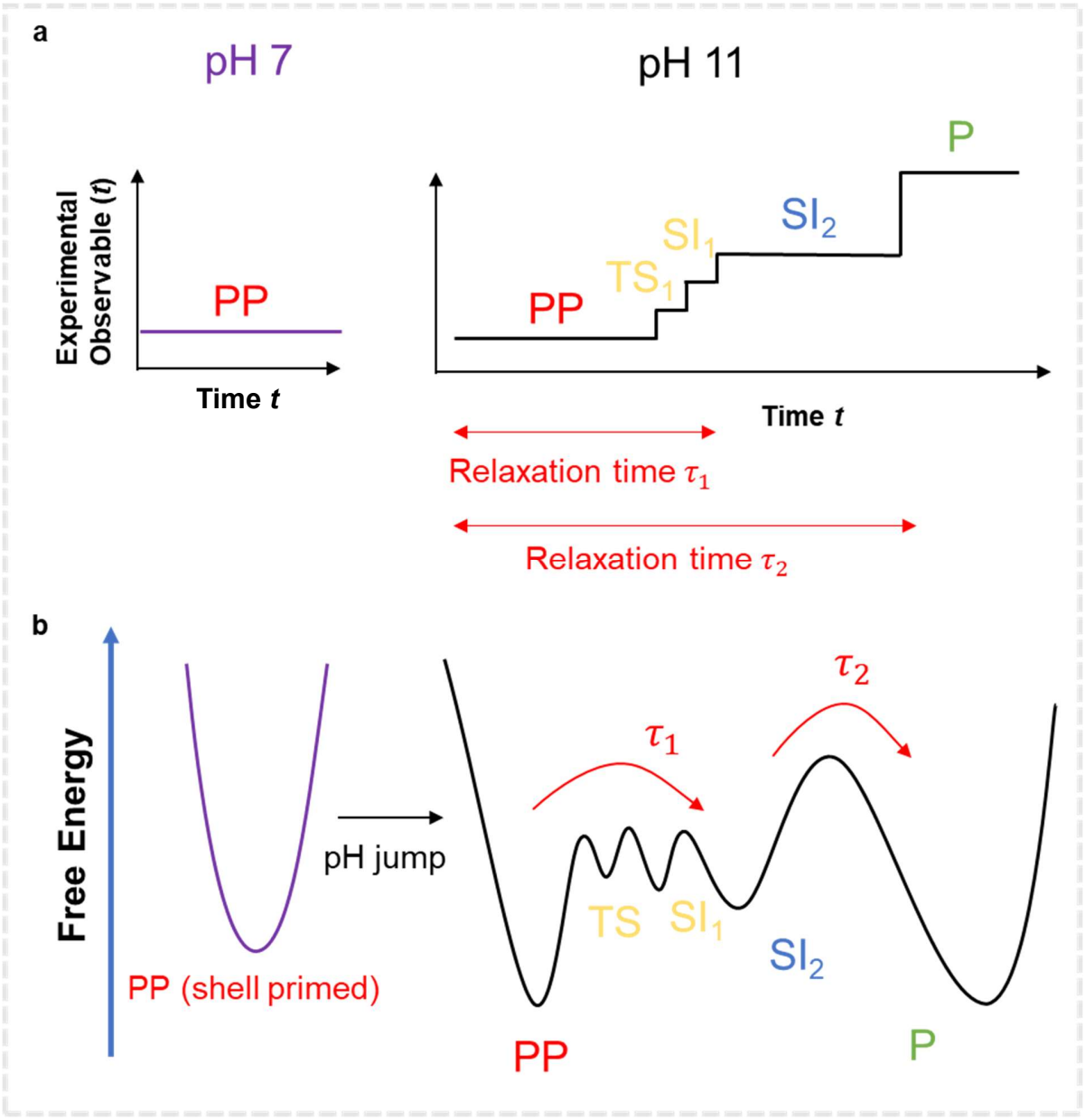


Figure 37: Energy landscape representation of all obtained results leading to a dynamic model. a Idealized experimental observable obtained by probing the complete transition of the Tc toxin. Under pH 7.5 conditions, the molecule remains in the PrePore state. After inducing a change to pH 11, the molecule performs a transition to the stable intermediate 2 (SI_2) via going over a transient intermediate 1 (TS) and another stable intermediate 1 (SI_1). After that, the pore is injected, so that the molecule reaches its final Pore state. **b** Corresponding energy landscape for the pH 7 environment showing only a static molecule (left) and the pH 11 environment, enabling the syringe like mechanism of the Tc toxin.

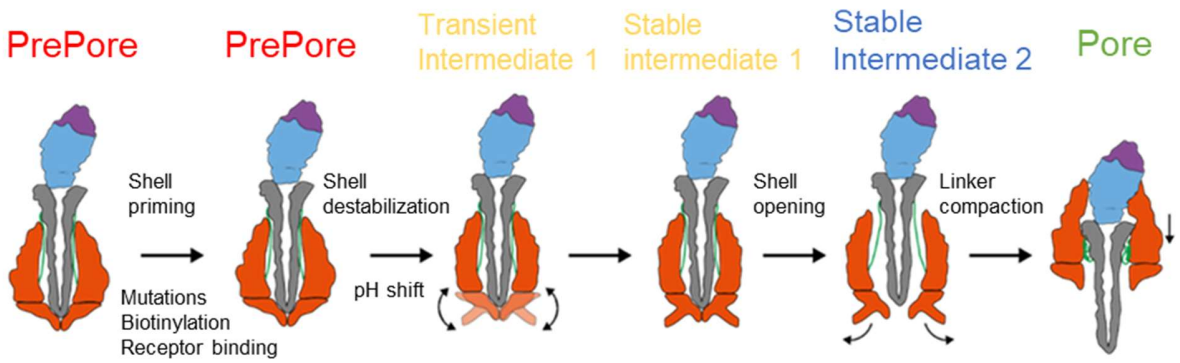


Figure 38: Schematic representation of the resolved states involved in the syringe like mechanism of the Tc toxin. Different factors can modulate the full pathway, which the molecule undergoes (green bars). After a pH shift, a flexibility of the surface of the shell is induced, resulting in a transient intermediate state. As next steps, the receptor binding domains B (on the bottom of the Tc toxin) move on the tip of the shell, resulting in a stable intermediate 1. Afterwards, the shell is opened, forming the stable intermediate 2. Next, the pore is injected and finally the Pore state is reached. Scheme created by D. Roderer and S. Raunser.

Since the results of this study originate from different type of experiments, data analysis and used models, a general overview of used samples, measured parameters and derived information can be found under Table 5 for investigation of state properties, and in Table 6 for investigation of kinetic properties of the Tc toxin.

It can be summarized that going from the PrePore to the Pore state of the molecule, it passes several involved states with rates that can be modulated using specific mutations or surface affecting actions like biotinylation. Using a mutant (K1197W) for which $k_1 = k_2 = 0$ (Table 6), we resolved a transient intermediate state TI-1, which shows a high flexibility of the surface of the protein. After a few minutes, a stable intermediate SI-1 is reached, showing only a very minor opening of the shell. Using the wild type, an opening of the shell could be observed, leading to a stable intermediate state SI-2. After longer waiting times, the Tc toxin reaches its final state, which is the Pore state, where the channel is ejected.

Studying the transition times of the conformational change from PrePore to SI-2 and from PrePore to Pore confirmed these findings (see Figure 37a). The PrePore to SI-2 transitions showed a broad distribution in the transition times, since they need to pass several states before reaching the SI-2 state. Connecting this with the previous results, we could resolve two of them, which is TI-1 and SI-1. We denote that it cannot be excluded that other states are involved. Monitoring the transition from the PrePore to the Pore state, we found that only a single step is performed using samples that can only sense the injection of the pore. From this we deduce that going from the SI-2 state to the final Pore state, no stable intermediate states are populated.

Combining these results, a kinetic pathway of the mechanism of the pore formation of the Tc toxin can be drawn (see Figure 37b and Figure 38). Here, multiple steps could be resolved and connected with each other answering the initial question of a kinetic model for the PrePore to Pore conformational change and mechanism (1) and if there are stable or transient intermediate states involved (2). On top, we answered question (3), that a perturbance of the shell via mutations or biotinylation leads to a dramatic change of the observed kinetic rates.

4.5 Conclusion

As presented, smFRET measurements and other fluorescence spectroscopy techniques can answer a variety of aspects on biomolecules and their dynamical behavior. In this chapter, I find that biomolecules are very sensitive towards small changes and their surrounding conditions.

For the HJ I resolved that under addition of Mg^{2+} ions, no stable planar cross state could be observed. This was achieved by using a new experimental approach, which is the OFA. As it could be seen, the OFA enhances signals and observation times, without perturbing the natural behavior of the investigated sample. As an outcome, I could follow the HJ in “real time” and could see it changing its state using time steps of only 50 µs. This method of resolving biomolecular dynamics has great potential since it is easy to communicate and does not rely on any deeper analysis.

In the case of the hGBP1 a large influence of the farnesylation was found. A direct comparison to the non-farnesylated hGBP1 showed that the farnesyl moiety works as an anchor fixing the molecule to a static state in the first place. Only after dimerization, dynamics are triggered. These dynamics appeared to be the same as in the non-farnesylated monomeric state. Next, the farnesylated hGBP1 only slowly populates the extended dimer, which is therefore only a short intermediate state on the way to its oligomerization, which is an exclusive ability of the farnesylated hGBP1. I find structures showing a good agreement to previously assumed models of the oligomeric state from literature.

For the Tc toxin it was possible to resolve its detailed pathway while it performs its syringe like mechanism, which is triggered by a pH shift. It could be seen that multiple steps are involved in this process. The experiments used not only freely diffusive samples, but also immobilized ones resolving transition times no other technique could reveal. We also demonstrated the possibility to modify the time scales on which the molecule acts dramatically by mutations or biotinylation. Combining these results leads to a consistent picture of the molecule.

5 Summary

In this thesis, fluorescence spectroscopy techniques were tested, further improved in terms of their reliability and used to resolve structures and dynamical behaviors of various biomolecules.

In the first main chapter it could be seen that using high precision single-molecule MFD it is possible to reliably measure the structure and dynamics of a protein. A worldwide benchmarking study showed high agreement in the obtained results, even though proteins are a complex and sensitive system. It was shown that smFRET measurements can sense structural conformational changes based upon ligand binding (MalE) and furthermore resolve complex multi-state dynamics (U2AF65) by using a consistent analysis over multiple laboratories. However, the agreement in the obtained data and analysis can be even further increased. Since user bias and an uncertainty in the estimation of correction parameters was observed, we suggest a uniform, more robust calibration workflow in a follow up study. Here, a protocol is developed with step-by-step instructions on how to accurately analyze single-molecule measurements. On top, we present a new way of determining the most critical parameter in calibration, which is the estimation of the spectrally different detection efficiencies. This is achieved by using a set of free dyes and connecting the obtained signals to the transmission spectra of every optical component of the setup. After an optimization process one obtains the desired detection efficiency ratio, which will be applicable to many different dye pairs used to measure FRET efficiencies. As a last point for chapter 3, a new idea is presented based on the photophysical effect of solvent relaxation. Since the dyes move due to their flexible linker they can access the surface of the protein as well as being away from it. Using different specific detection wavelengths in the experiment, an estimation of the ratio of dyes on the surface or away from it could be done. Since dyes close to the surface can experience quenching and sticking, emission from these dyes can be partly filtered out. This will be beneficial in the event of a dye pair showing high anisotropy values as was the case in the presented benchmark study. Here, one could try to “rescue” these samples using the described method, to achieve an accurate data analysis.

In the second main chapter biomolecules are investigated which vary in many aspects: They have masses ranging from 70 kDa to 1700 kDa, and perform different functions like genetic recombination (HJ), human immune response (hGBP1) and attacking cells via perforating the membrane (Tc toxin). To gain insights into their function, their structure und dynamical behavior was explored. It was possible by using an OFA to watch the transition of the Holliday Junction going from one stacked state to the other in real time (sub-ms-resolution). Such FRET based experiments are a “dream of every biochemist” [1] since they allow to track the molecule’s mechanism while it performs its complex function. For the HJ it was seen that the suspected cross shape state is only a fast decaying transition state. As a next target, the dynamics and structure of hGBP1 was investigated. We refined the structure of the non-farnesylated hGBP1 in greater detail using additional FRET pairs. The focus of this study is however the comparison of the non- to the farnesylated hGBP1. Here it was found that the farnesyl moiety works as an anchor prohibiting the dynamics in the monomeric state that has been seen for the non-farnesylated hGBP1. After a dimer is formed and a substrate is bound, the same dynamics reoccur. This allows the molecule to follow the pathway from the monomer, via the dimer to the higher order oligomer in order to fulfil its biological function. While the non-farnesylated hGBP1 does not pass the elongated dimer state on the pathway, the oligomerization is exclusively performed by the farnesylated hGBP1. As a final target the syringe like mechanism of the Tc toxin was probed. After combining different experimental approaches, we found that the shell opening of the Tc toxin is a complex conformational transition involving stable and transient intermediate states. Contrary, the injection of the shell was seen to be a consecutive but discrete process. To resolve the

conformational transitions path times, we performed TIRF measurements and monitored the Tc toxin live in action. Here our previous findings were confirmed and we found a broad distribution of transition path times for the shell opening mechanism, whereas only a step wise change in the experimental observable was seen while probing the pore injection. With these findings we answered the question how the Tc toxin performs this syringe like mechanism while it is perforating the membrane of a cell.

This thesis has several contributions to current questions of science which are related to both methodological improvements and to the insights into the structure, dynamics and function of biomolecules. By combining the gained insights from the global FRET study with our presented uniform workflow, a robust way of using smFRET experiments to study biomolecules for every laboratory in the world using or planning to do these is established. With this technique, we could follow a process in real time which has a major role in genetic recombination, hence it is a basic process crucial for life. Furthermore, we explored how a protein performs a conformational pathway from its monomeric to oligomeric state to fulfil its biological function in the immune defense of a human. Finally, we resolved a syringe like mechanism of a Tc toxin, which will help in using this molecule as a biopesticide or in medical related science.

References

1. Henzler-Wildman, K. and D. Kern, *Dynamic Personalities of Proteins.* Nature, 2007. **450**(7172): p. 964-972.
2. Kendrew, J.C., et al., *3-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis.* Nature, 1958. **181**(4610): p. 662-666.
3. Kupitz, C., et al., *Serial Time-Resolved Crystallography of Photosystem Ii Using a Femtosecond X-Ray Laser.* Nature, 2014. **513**(7517): p. 261-+.
4. Frueh, D.P., et al., *Nmr Methods for Structural Studies of Large Monomeric and Multimeric Proteins.* Current Opinion in Structural Biology, 2013. **23**(5): p. 734-739.
5. Anthis, N.J. and G.M. Clore, *Visualizing Transient Dark States by Nmr Spectroscopy.* Quarterly Reviews of Biophysics, 2015. **48**(1): p. 35-116.
6. Palmer, A.G., *Nmr Characterization of the Dynamics of Biomacromolecules.* Chemical Reviews, 2004. **104**(8): p. 3623-3640.
7. Ravera, E., et al., *Insights into Domain–Domain Motions in Proteins and Rna from Solution Nmr.* Accounts of Chemical Research, 2014. **47**(10): p. 3118-3126.
8. Sekhar, A. and L.E. Kay, *An Nmr View of Protein Dynamics in Health and Disease.* Annual Review of Biophysics, 2019. **48**(1): p. 297-319.
9. Feng, X.S., et al., *A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-Em.* Structure, 2017. **25**(4): p. 663-+.
10. Jeschke, G., *Deer Distance Measurements on Proteins.* Annual Review of Physical Chemistry, 2012. **63**(1): p. 419-446.
11. Jumper, J., et al., *Highly Accurate Protein Structure Prediction with Alphafold.* Nature, 2021. **596**(7873): p. 583.
12. de Oliveira, S. and C.A.-O. Deane, *Co-Evolution Techniques Are Reshaping the Way We Do Structural Bioinformatics.* (2046-1402 (Print)).
13. Orozco, M., *A Theoretical View of Protein Dynamics.* Chem. Soc. Rev., 2014. **43**(14): p. 5051-5066.
14. Orrit, M. and J. Bernard, *Single Pentacene Molecules Detected by Fluorescence Excitation in a Para-Terphenyl Crystal.* Physical Review Letters, 1990. **65**(21): p. 2716-2719.
15. Ha, T., et al., *Probing the Interaction between Two Single Molecules: Fluorescence Resonance Energy Transfer between a Single Donor and a Single Acceptor.* Proceedings of the National Academy of Sciences, 1996. **93**(13): p. 6264-6268.
16. Hellenkamp, B., et al., *Precision and Accuracy of Single-Molecule Fret Measurements-a Multi-Laboratory Benchmark Study (Vol 15, Pg 984, 2018).* Nature Methods, 2018. **15**(11): p. 984-984.
17. Dimura, M., et al., *Automated and Optimally Fret-Assisted Structural Modeling.* Nature Communications, 2020. **11**(1).

18. Kühnemuth, R. and C.A.M. Seidel, *Principles of Single Molecule Multiparameter Fluorescence Spectroscopy.* Single Molecules, 2001. **2**(4): p. 251-254.
19. Kudryavtsev, V., et al., *Combining Mfd and Pie for Accurate Single-Pair Forster Resonance Energy Transfer Measurements.* Chemphyschem, 2012. **13**(4): p. 1060-1078.
20. Levinthal, C., *Are There Pathways for Protein Folding.* Journal De Chimie Physique Et De Physico-Chimie Biologique, 1968. **65**(1): p. 44.
21. Seidel, C., *Biophysikalische Grundlagen Der Biomolekularen Strukturbildung.* Skript, 2021.
22. Schultz, C.P., *Illuminating Folding Intermediates.* Nature Structural Biology, 2000. **7**(1): p. 7-10.
23. Dill, K.A., *Theory for the Folding and Stability of Globular-Proteins.* Biochemistry, 1985. **24**(6): p. 1501-1509.
24. Austin, R.H., et al., *Dynamics of Ligand-Binding to Myoglobin.* Biochemistry, 1975. **14**(24): p. 5355-5373.
25. Frauenfelder, H., S.G. Sligar, and P.G. Wolynes, *The Energy Landscapes and Motions of Proteins.* Science, 1991. **254**(5038): p. 1598-1603.
26. Ansari, A., et al., *Protein States and Protein Quakes.* Proceedings of the National Academy of Sciences of the United States of America, 1985. **82**(15): p. 5000-5004.
27. Chung, H.S., et al., *Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times.* Science, 2012. **335**(6071): p. 981-984.
28. Goodsell, D.S., *A Look inside the Living Cell.* American Scientist, 1992. **80**(5): p. 457-465.
29. Choi, I.G. and S.H. Kim, *Evolution of Protein Structural Classes and Protein Sequence Families.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(38): p. 14056-14061.
30. Lerner, E., et al., *Fret-Based Dynamic Structural Biology: Challenges, Perspectives and an Appeal for Open-Science Practices.* Elife, 2021. **10**.
31. Budde, J.-H., et al. *Fret Nanoscopy Enables Seamless Imaging of Molecular Assemblies with Sub-Nanometer Resolution*. 2021. arXiv:2108.00024.
32. Kalinin, S., et al., *Detection of Structural Dynamics by Fret: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States.* Journal of Physical Chemistry B, 2010. **114**(23): p. 7983-7995.
33. Barth, A., et al., *Unraveling Multi-State Molecular Dynamics in Single-Molecule Fret Experiments. I. Theory of Fret-Lines.* Journal of Chemical Physics, 2022. **156**(14).
34. Felekyan, S., et al., *Filtered Fcs: Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules.* Chemphyschem, 2012. **13**(4): p. 1036-1053.

35. Sanchez-Rico, C., et al., *Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by Spfret and Nmr Spectroscopy.* Chemistry-a European Journal, 2017. **23**(57): p. 14267-14277.
36. von Voithenberg, L.V., et al., *Recognition of the 3 ' Splice Site Rna by the U2af Heterodimer Involves a Dynamic Population Shift.* Proceedings of the National Academy of Sciences of the United States of America, 2016. **113**(46): p. E7169-E7175.
37. Torella, J.P., et al., *Identifying Molecular Dynamics in Single-Molecule Fret Experiments with Burst Variance Analysis.* Biophysical Journal, 2011. **100**(6): p. 1568-1577.
38. Gopich, I.V. and A. Szabo, *Theory of the Energy Transfer Efficiency and Fluorescence Lifetime Distribution in Single-Molecule Fret.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(20): p. 7747-7752.
39. Zaccai, G., *Biochemistry - How Soft Is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering.* Science, 2000. **288**(5471): p. 1604-1607.
40. Fries, J.R., et al., *Quantitative Identification of Different Single Molecules by Selective Time-Resolved Confocal Fluorescence Spectroscopy.* Journal of Physical Chemistry A, 1998. **102**(33): p. 6601-6613.
41. Tomov, T.E., et al., *Disentangling Subpopulations in Single-Molecule Fret and Alex Experiments with Photon Distribution Analysis.* Biophysical Journal, 2012. **102**(5): p. 1163-1173.
42. Coullomb, A., et al., *Quanti-Fret: A Framework for Quantitative Fret Measurements in Living Cells.* Scientific Reports, 2020. **10**(1).
43. Lee, N.K., et al., *Accurate Fret Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation.* Biophysical Journal, 2005. **88**(4): p. 2939-2953.
44. Peulen, T.O., O. Opanasyuk, and C.A.M. Seidel, *Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved Fret Measurements of Labeled Macromolecules Accurately.* Journal of Physical Chemistry B, 2017. **121**(35): p. 8211-8241.
45. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy.* Srpinger, 2006.
46. Chen, H.M., et al., *Mechanisms of Quenching of Alexa Fluorophores by Natural Amino Acids.* Journal of the American Chemical Society, 2010. **132**(21): p. 7244.
47. Mockel, C., et al., *Integrated Nmr, Fluorescence, and Molecular Dynamics Benchmark Study of Protein Mechanics and Hydrodynamics.* Journal of Physical Chemistry B, 2019. **123**(7): p. 1453-1480.
48. Kalinin, S., et al., *A Toolkit and Benchmark Study for Fret-Restrained High-Precision Structural Modeling.* Nature Methods, 2012. **9**(12): p. 1218-U129.
49. Holliday, R., *Mechanism for Gene Conversion in Fungi.* Genetical Research, 1964. **5**(2): p. 282-&.

50. Broker, T.R. and I.R. Lehman, *Branched DNA Molecules - Intermediates in T4 Recombination.* Journal of Molecular Biology, 1971. **60**(1): p. 131.
51. Sigal, N. and B. Alberts, *Genetic Recombination - Nature of a Crossed Strand-Exchange between 2 Homologous DNA Molecules.* Journal of Molecular Biology, 1972. **71**(3): p. 789.
52. Zettl, T., et al., *The Structural Ensemble of a Holliday Junction Determined by X-Ray Scattering Interference.* Nucleic Acids Research, 2020. **48**(14): p. 8090-8098.
53. Lilley, D.M.J., *Structures of Helical Junctions in Nucleic Acids.* Quarterly Reviews of Biophysics, 2000. **33**(2): p. 109-159.
54. McKinney, S.A., et al., *Spontaneous and Protein-Induced Structural Dynamics of Individual Holliday Junctions.* Biophysical Journal, 2003. **84**(2): p. 303a-304a.
55. Valeri, A., *Fluorescence Resonance Energy Transfer between Multiple Chromophores Studied by Single Molecule Spectroscopy.* Dissertation, Heinrich-Heine University, Düsseldorf, 2009.
56. Valeri, A., et al., *Single-Molecule Fret Detects Intermediates and Fast Dynamics of DNA Holliday Junctions.* unpublished work.
57. Clegg, R.M., A.I.H. Murchie, and D.M.J. Lilley, *The Solution Structure of the 4-Way DNA Junction at Low-Salt Conditions - a Fluorescence Resonance Energy-Transfer Analysis.* Biophysical Journal, 1994. **66**(1): p. 99-109.
58. Gietl, A., et al., *DNA Origami as Biocompatible Surface to Match Single-Molecule and Ensemble Experiments.* Nucleic Acids Research, 2012. **40**(14).
59. Lee, K.G., et al., *A Planar Dielectric Antenna for Directional Single-Photon Emission and near-Unity Collection Efficiency.* Nature Photonics, 2011. **5**(3): p. 166-169.
60. Wang, Q., et al., *Probing Single Biomolecules in Solution Using the Anti-Brownian Electrokinetic (Abel) Trap.* Accounts of Chemical Research, 2012. **45**(11): p. 1955-1964.
61. Inostroza, L.M., *Planar Dielectric Antennas for Applications in Bio-Physics.* Dissertation, Friedrich-Alexander-Universität, Erlangen, 2022.
62. McKinney, S.A., C. Joo, and T. Ha, *Analysis of Single-Molecule Fret Trajectories Using Hidden Markov Modeling.* Biophysical Journal, 2006. **91**(5): p. 1941-1951.
63. Cohen, A.E., et al., *The Anti-Brownian Electrophoretic Trap (Abel Trap): Fabrication and Software*, in *Imaging, Manipulation, and Analysis of Biomolecules and Cells: Fundamentals and Applications III*. 2005.
64. Wilson, H. and Q. Wang, *Abel-Fret: Tether-Free Single-Molecule Fret with Hydrodynamic Profiling.* Nature Methods, 2021. **18**(7): p. 816.
65. Koenderink, A.F., *Single-Photon Nanoantennas.* Acs Photonics, 2017. **4**(4): p. 710-722.

66. Shrivastav, A.M., U. Cvelbar, and I. Abdulhalim, *A Comprehensive Review on Plasmonic-Based Biosensors Used in Viral Diagnostics.* Communications Biology, 2021. **4**(1).
67. Glembockyte, V., et al., *DNA Origami Nanoantennas for Fluorescence Enhancement.* Accounts of Chemical Research, 2021. **54**(17): p. 3338-3348.
68. Chung, H.S., J.M. Louis, and W.A. Eaton, *Experimental Determination of Upper Bound for Transition Path Times in Protein Folding from Single-Molecule Photon-by-Photon Trajectories.* Proceedings of the National Academy of Sciences, 2009. **106**(29): p. 11837-11844.
69. Chung, H.S., *Transition Path Times Measured by Single-Molecule Spectroscopy.* Journal of Molecular Biology, 2018. **430**(4): p. 409-423.
70. Kim, J.Y. and H.S. Chung, *Disordered Proteins Follow Diverse Transition Paths as They Fold and Bind to a Partner.* Science, 2020. **368**(6496): p. 1253.
71. Praefcke, G.J.K. and H.T. McMahon, *The Dynamin Superfamily: Universal Membrane Tubulation and Fission Molecules?* Nature Reviews Molecular Cell Biology, 2004. **5**(2): p. 133-147.
72. MacMicking, J.D., *Interferon-Inducible Effector Mechanisms in Cell-Autonomous Immunity.* Nature Reviews Immunology, 2012. **12**(5): p. 367-382.
73. Tretina, K., et al., *Interferon-Induced Guanylate-Binding Proteins: Guardians of Host Defense in Health and Disease.* Journal of Experimental Medicine, 2019. **216**(3): p. 482-500.
74. Praefcke, G.J.K., *Regulation of Innate Immune Functions by Guanylate-Binding Proteins.* International Journal of Medical Microbiology, 2018. **308**(1): p. 237-245.
75. Honkala, A.T., D. Tailor, and S.V. Malhotra, *Guanylate-Binding Protein 1: An Emerging Target in Inflammation and Cancer.* Frontiers in Immunology, 2020. **10**.
76. Ghosh, P., et al., *Structural Mechanism for Guanylate-Binding Proteins (Gbps) Targeting by the Shigella E3 Ligase Ipah9.8.* PLOS Pathogens, 2019. **15**(6).
77. Shydlovskyi, S., et al., *Nucleotide-Dependent Farnesyl Switch Orchestrates Polymerization and Membrane Binding of Human Guanylate-Binding Protein 1.* Proceedings of the National Academy of Sciences, 2017. **114**(28).
78. Ince, S., et al., *The Human Guanylate-Binding Proteins Hgbp-1 and Hgbp-5 Cycle between Monomers and Dimers Only.* The FEBS Journal, 2017. **284**(14): p. 2284-2301.
79. Ghosh, A., et al., *How Guanylate-Binding Proteins Achieve Assembly-Stimulated Processive Cleavage of Gtp to Gmp.* Nature, 2006. **440**(7080): p. 101-104.
80. Sistemich, L., et al., *The Molecular Mechanism of Polymer Formation of Farnesylated Human Guanylate-Binding Protein 1.* Journal of Molecular Biology, 2020. **432**(7): p. 2164-2185.

81. Sistemich, L., et al., *Structural Requirements for Membrane Binding of Human Guanylate-Binding Protein 1.* The FEBS Journal, 2021. **288**(13): p. 4098-4114.
82. Fernandez-Fuentes, N., et al., *Exploring Protein Dynamics Space: The Dynasome as the Missing Link between Protein Structure and Function.* PLoS ONE, 2012. **7**(5).
83. Mittag, T., L.E. Kay, and J.D. Forman-Kay, *Protein Dynamics and Conformational Disorder in Molecular Recognition.* Journal of Molecular Recognition, 2009.
84. Peulen, T.-O., et al. *Integrative Dynamic Structural Biology Unveils Conformers Essential for the Oligomerization of a Large Gtpase*. 2020. arXiv:2004.04229.
85. Vöpel, T., et al., *Triphosphate Induced Dimerization of Human Guanylate Binding Protein 1 Involves Association of the C-Terminal Helices: A Joint Double Electron–Electron Resonance and Fret Study.* Biochemistry, 2014. **53**(28): p. 4590-4600.
86. Cui, W., et al., *Structural Basis for Gtp-Induced Dimerization and Antiviral Function of Guanylate-Binding Proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2021. **118**(15).
87. Zhu, S., et al., *Cryo-Et of a Human Gbp Coatomer Governing Cell-Autonomous Innate Immunity to Infection.* bioRxiv, 2021.
88. Roderer, D., et al., *Towards the Application of Tc Toxins as a Universal Protein Translocation System.* Nature Communications, 2019. **10**.
89. Roderer, D. and S. Raunser, *Tc Toxin Complexes: Assembly, Membrane Permeation, and Protein Translocation.* Annual Review of Microbiology, Vol 73, 2019. **73**: p. 247.
90. Shimamura, T., et al., *Molecular Basis of Alternating Access Membrane Transport by the Sodium-Hydantoin Transporter Mhp1.* Science, 2010. **328**(5977): p. 470-473.
91. Bowen, D., et al., *Insecticidal Toxins from the Bacterium Photorhabdus Luminescens.* Science, 1998. **280**(5372): p. 2129-2132.
92. Sergeant, M., et al., *Interactions of Insecticidal Toxin Gene Products from Xenorhabdus Nematophilus Pmf1296.* Applied and Environmental Microbiology, 2003. **69**(6): p. 3344-3349.
93. ffrench-Constant, R.H. and D.J. Bowen, *Novel Insecticidal Toxins from Nematode-Symbiotic Bacteria.* Cellular and Molecular Life Sciences, 2000. **57**(5): p. 828-833.
94. Gerrard, J., et al., *Human Infection with Photorhabdus Asymbiotica: An Emerging Bacterial Pathogen.* Microbes and Infection, 2004. **6**(2): p. 229-237.
95. Tennant, S.M., et al., *Homologues of Insecticidal Toxin Complex Genes in Yersinia Enterocolitica Biotype 1a and Their Contribution to Virulence.* Infection and Immunity, 2005. **73**(10): p. 6860-6867.

96. Waterfield, N., et al., *The Insect Toxin Complex of Yersinia.* Genus Yersinia: From Genomics to Function, 2007. **603**: p. 247-257.
97. ffrench-Constant, R.H., A. Dowling, and N.R. Waterfield, *Insecticidal Toxins from Photorhabdus Bacteria and Their Potential Use in Agriculture.* Toxicon, 2007. **49**(4): p. 436-451.
98. Pinheiro, V.B. and D.J. Ellar, *Expression and Insecticidal Activity of Yersinia Pseudotuberculosis and Photorhabdus Luminescens Toxin Complex Proteins.* Cellular Microbiology, 2007. **9**(10): p. 2372-2380.
99. Zahaf, N.I., et al., *Targeted Delivery of an Adp-Ribosylating Bacterial Toxin into Cancer Cells.* Scientific Reports, 2017. **7**.
100. Gatsogiannis, C., et al., *A Syringe-Like Injection Mechanism in Photorhabdus Luminescens Toxins.* Nature, 2013. **495**(7442): p. 520-523.
101. Meusch, D., et al., *Mechanism of Tc Toxin Action Revealed in Molecular Detail.* Nature, 2014. **508**(7494): p. 61.
102. Gatsogiannis, C., et al., *Membrane Insertion of a Tc Toxin in near-Atomic Detail.* Nature Structural & Molecular Biology, 2016. **23**(10): p. 884-890.
103. Roderer, D., et al., *Structure of a Tc Holotoxin Pore Provides Insights into the Translocation Mechanism.* Proceedings of the National Academy of Sciences of the United States of America, 2019. **116**(46): p. 23083-23090.
104. Roderer, D., et al., *Glycan-Dependent Cell Adhesion Mechanism of Tc Toxins.* Nature Communications, 2020. **11**(1).
105. Ortega, A., D. Amoros, and J.G. de la Torre, *Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models.* Biophysical Journal, 2011. **101**(4): p. 892-898.
106. Widengren, J., et al., *Single-Molecule Detection and Identification of Multiple Species by Multiparameter Fluorescence Detection.* Analytical Chemistry, 2006. **78**(6): p. 2039-2050.
107. Kalinin, S., et al., *Probability Distribution Analysis of Single-Molecule Fluorescence Anisotropy and Resonance Energy Transfer.* Journal of Physical Chemistry B, 2007. **111**(34): p. 10253-10262.
108. Preus, S., et al., *Isms: Single-Molecule Fret Microscopy Software.* Nature Methods, 2015. **12**(7): p. 593-594.
109. Devauges, V., et al., *Homodimerization of Amyloid Precursor Protein at the Plasma Membrane: A Homofret Study by Time-Resolved Fluorescence Anisotropy Imaging.* Plos One, 2012. **7**(9).
110. Strohl, F., et al., *Total Internal Reflection Fluorescence Anisotropy Imaging Microscopy: Setup, Calibration, and Data Processing for Protein Polymerization Measurements in Living Cells.* Methods and Applications in Fluorescence, 2018. **6**(1).
111. Corp., C.T. *Spectra Viewer*. 2021; Available from: https://www.chroma.com/spectra-viewer.
112. Hellenkamp, B., et al., *Precision and Accuracy of Single-Molecule Fret Measurements-a Multi-Laboratory Benchmark Study.* Nature Methods, 2018. **15**(9): p. 669-+.

113. Sindbert, S., et al., *Accurate Distance Determination of Nucleic Acids Via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity.* Journal of the American Chemical Society, 2011. **133**(8): p. 2463-2480.
114. Antonik, M., et al., *Separating Structural Heterogeneities from Stochastic Variations in Fluorescence Resonance Energy Transfer Distributions Via Photon Distribution Analysis.* Journal of Physical Chemistry B, 2006. **110**(13): p. 6970-6978.
115. Lubeseder-Martellato, C., et al., *Guanylate-Binding Protein-1 Expression Is Selectively Induced by Inflammatory Cytokines and Is an Activation Marker of Endothelial Cells During Inflammatory Diseases.* The American Journal of Pathology, 2002. **161**(5): p. 1749-1759.
116. Pilla-Moffett, D., et al., *Interferon-Inducible Gtpases in Host Resistance, Inflammation and Disease.* Journal of Molecular Biology, 2016. **428**(17): p. 3495-3513.
117. Xavier, A., et al., *Hgbp1 Coordinates Chlamydia Restriction and Inflammasome Activation through Sequential Gtp Hydrolysis.* Cell Reports, 2020. **31**(7).
118. Fisch, D., et al., *Human Gbp 1 Is a Microbe-Specific Gatekeeper of Macrophage Apoptosis and Pyroptosis.* The EMBO Journal, 2019. **38**(13).
119. Tripal, P., et al., *Unique Features of Different Members of the Human Guanylate-Binding Protein Family.* Journal of Interferon & Cytokine Research, 2007. **27**(1): p. 44-52.
120. Prakash, B., et al., *Structure of Human Guanylate-Binding Protein 1 Representing a Unique Class of Gtp-Binding Proteins.* Nature, 2000. **403**(6769): p. 567-571.
121. Lorenz, C., et al., *Farnesylation of Human Guanylate-Binding Protein 1 as Safety Mechanism Preventing Structural Rearrangements and Uninduced Dimerization.* The FEBS Journal, 2019. **287**(3): p. 496-514.
122. Ahmed, A. and H. Gohlke, *Multiscale Modeling of Macromolecular Conformational Changes Combining Concepts from Rigidity and Elastic Network Theory.* Proteins-Structure Function and Bioinformatics, 2006. **63**(4): p. 1038-1051.
123. Vopel, T., et al., *Triphosphate Induced Dimerization of Human Guanylate Binding Protein 1 Involves Association of the C-Terminal Helices: A Joint Double Electron-Electron Resonance and Fret Study.* Biochemistry, 2014. **53**(28): p. 4590-4600.
124. Sisamakis, E., et al., *Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection.* Methods in Enzymology, Vol 475: Single Molecule Tools, Pt B, 2010. **475**: p. 455-514.
125. Shydlovskyi, S., et al., *Nucleotide-Dependent Farnesyl Switch Orchestrates Polymerization and Membrane Binding of Human Guanylate-Binding Protein 1.* Proceedings of the National Academy of Sciences of the United States of America, 2017. **114**(28): p. E5559-E5568.

126. Sistemich, L. and C. Herrmann, *Purification of and Characterization of Its Polymerization and Membrane Binding*, in *Dynamin Superfamily Gtpases*. 2020. p. 67-81.
127. Hendrix, J. and D.C. Lamb, *Pulsed Interleaved Excitation: Principles and Applications.* Fluorescence Fluctuation Spectroscopy (Ffs), Pt A, 2013. **518**: p. 205-243.
128. Dimura, M., et al., *Quantitative Fret Studies and Integrative Modeling Unravel the Structure and Dynamics of Biomolecular Systems.* Current Opinion in Structural Biology, 2016. **40**: p. 163-185.
129. Wehner, M., S. Kunzelmann, and C. Herrmann, *The Guanine Cap of Human Guanylate-Binding Protein 1 Is Responsible for Dimerization and Self-Activation of Gtp Hydrolysis.* FEBS Journal, 2012. **279**(2): p. 203-210.
130. Hengstenberg, C.S., *Structural Dynamics and Implications for Dimer Formation of Human Guanylate-Binding Protein 1*. 2015.

Contributions

Publikationen:

1. Agam, G.[#] , Gebhardt, C.[#] , Popara, M.[#] , Mächtel, R., **Folz, J.**, Ambrose, B., Chamachi, N., Chung S. Y., Craggs, T. D., de Boer, M., Grohmann, D., Ha, T., Hartmann, A., Hendrix J., Hirschfeld, V, Hübner, C. G., Hugel T., Kammerer D., Kang H.-S., Kapanidis, A. N., Krainer, G., Kram, K., Lemke, E., Lerner, E., Margeat, E., Martens, K., Michaelis, J., Mitra J., Moya Muñoz G. G., Quast, R., Robb N. B., Sattler, M., Schlierf, M. Schneider, J., Schröder, T., Sefer, A., Tan, P. T., Thurn J., Tinnefeld, P., van Noort, J., Weiss, S., Wendler, N., Zijlstra, N., Barth, A.*, Seidel, C. A. M.*, Lamb, D. C.*, Cordes, T.*; *Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins*. Nat. Methods 20, 523–535 (2023). doi.org/10.1038/s41592-023-01807-0. (Chapter 3.2 and 7.1)
2. Peulen, T.-O., Hengstenberg, C.S., Biehl, R., Dimura, M., Lorenz, C., Valeri, A., **Folz, J.**, Hanke, C. A., Ince, S., Vöpel, T., Faragó, B., Gohlke, H., Klare, J.P., Stadler, A., Seidel, C.A.M.*, Herrmann, C.*; *Integrative dynamic structural biology unveils conformers essential for the oligomerization of a large GTPase*. *eLife* **12**:e79565. https://doi.org/10.7554/eLife.79565 (p. 52 et seq.)

Preprints:

3. Morales-Inostroza, L.[#], **Folz, J.** [#], Kühnemuth, R., Felekyan, S. Wieser, F.-F., Seidel, C. A. M.*, Götzinger, S., Sandoghdar, V.*; *An optofluidic antenna for enhancing the sensitivity of single-emitter measurements*, 27 January 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-2497202/v1]. doi.org/10.21203/rs.3.rs2497202/v1. In Revision bei Nature Commun. (Chapter 4.2 and 7.2)
4. van der Voort, N.T.M., Budde, J.-H., Felekyan, S., **Folz, J.**, Kühnemuth, R., Lauterjung, P., Köhler, M., Schönle, A., Sindram, J., Otten, M., Karg, M., Herrmann, C., Barth, A. and Seidel, C. A. M.; *FRET nanoscopy enables seamless imaging of molecular assemblies with sub-nanometer resolution*. arXiv:2108.00024v2 (p. 52 et seq.)

Zur Veröffentlichung vorbereitete Manuskripte:

5. **Folz J.[#]**, Lauterjung, P.[#], Seidel, C.A.M., Herrmann C, *Intramolecular dynamics and function of hGBP1 are controlled by farnesylation.* (Chapter 4.3)
6. Ng'ang'a, P. N.[#], Kucher, S.[#], **Folz, J.[#]**, Roderer, D.[#], Xu, Y., Sitsel, O., Belyy, A., Prumbaum, D., Kühnemuth, R., Dong, M., Seidel, C.A.M.*, Bordignon, E.*, Raunser, S.*; *Kinetics of the syringe-like injection mechanism of Tc toxins*. (Chapter 4.4)
7. Valeri, A., Felekyan, S., Kalinin, S., Richert. M., Marawske, S., Schweinberger E., **Folz, J.**, Kühnemuth R., Rech I.,Gulinatti, A., Seidel, C. A. M.; *Single-molecule FRET detects intermediates and fast dynamics of a DNA Holliday junction*. (p. 42 et seq.)

Contribution to conferences

Talks

J. Folz, S. Felekyan, R. Kühnemuth and C.A.M. Seidel
"FRET measurements resolving structures, dynamics and transitions of various biomolecules"
Murnau Conference, Murnau 2022

J. Folz, S. Felekyan, R. Kühnemuth and C.A.M. Seidel
"FRET measurements resolving structures, dynamics and transitions of various biomolecules"
GBM Meeting, Düsseldorf 2022

J. Folz, T.-O. Peulen, O. Opanasyuk and C.A.M. Seidel
"Accuracy in FRET measurements concerning technical and methodical aspects"
Single-molecule Workshop, PicoQuant, Berlin 2018

J. Folz, A. Larbig and C.A.M. Seidel
"Dipolar Coupling as viewed from FRET spectroscopic techniques"
COST International Scientific Meeting, Lissabon, 2018

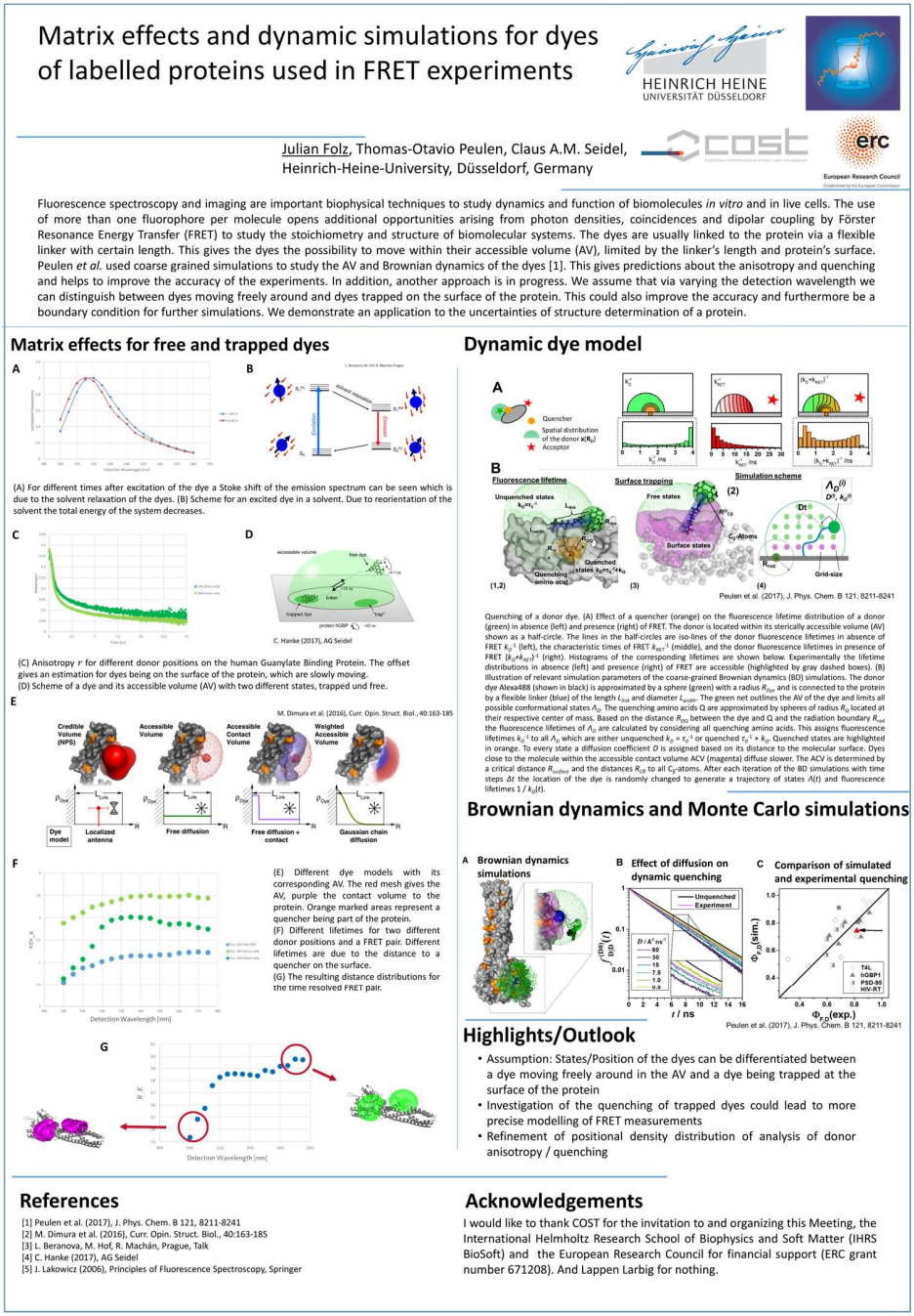
J. Folz, S. Felekyan, R. Kühnemuth and C.A.M. Seidel
"FRET measurements resolving structures, dynamics and transitions of various biomolecules"
Single-molecule Workshop, PicoQuant, Berlin 2021

Posters

J. Folz, T.-O. Peulen and C.A.M. Seidel

"Matrix effects and dynamic simulations for dyes of labelled proteins used in FRET experiments"

4th International Scientific Meeting and 5th MC Meeting of COST Action CM1306, Zagreb 2018



J. Folz, T.-O. Peulen, O. Opanasyuk and C.A.M. Seidel

"Accuracy in FRET measurements concerning technical and methodical aspects"

7th Murnau Conference on Structural Biology, Murnau, 2018

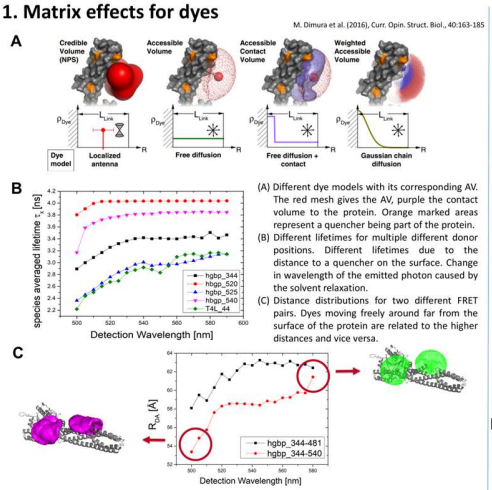
Accuracy in FRET measurements concerning technical and methodical aspects

J. Folz, T.-O. Peulen, O. Opanasyuk and C.A.M. Seidel
Heinrich-Heine-University, Düsseldorf

Fluorescence spectroscopy and imaging are important biophysical techniques to study dynamics and function of biomolecules *in vitro* and in live cells. The use of more than one fluorophore per molecule opens additional opportunities arising from photon densities, coincidences and dipolar coupling by Förster Resonance Energy Transfer (FRET) to study the stoichiometry and structure of biomolecular systems. Before performing FRET measurements one needs to consider which data should be acquired and what information is contained in the data with what accuracy. We compared data of fluorescence lifetime and fluorescence intensity experiments for the sake of FRET based-structural integrative modelling though dye models. Different models of the fluorescent dyes used to model observables of FRET experiments will be introduced and discussed concerning their accuracy. Exploiting the solvatochromism of fluorescent dyes in time-resolved fluorescence experiments, we probe experimentally fully solvated dyes and dyes bound to protein surfaces. Such data will be presented for a network of labeling sites of a large GTPase and may be utilized for refined coarse-grained dye models for high-speed integrative structural modelling at high accuracy.

Accuracy in FRET measurements

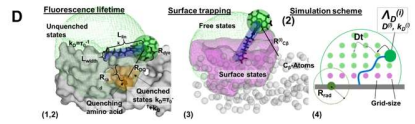
1. Matrix effects for dyes

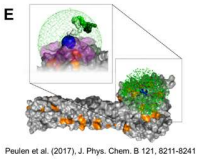


(A) Different dye models with its corresponding AV. The red mesh gives the AV, purple the contact volume to the protein. Orange marked areas represent a quencher being part of the protein.
(B) Different lifetimes for multiple different donor positions. Different lifetimes due to the distance to a quencher on the surface. Change in wavelength of the emitted photon caused by the solvent relaxation.
(C) Distance distributions for two different FRET pairs. Dyes moving freely around far from the surface of the protein are related to the higher distances and vice versa.

Matrix effects for dyes like quenching and solvatochrism could be used for estimating the approximate position of the dye in the accessible volume

2. Dynamic dye model

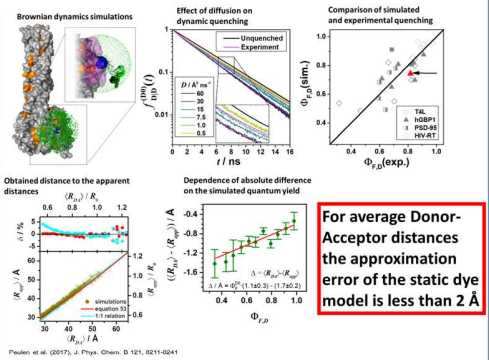




(D) Illustration of relevant simulation parameters of the coarse-grained Brownian dynamics (BD) simulations.
(E) Example of a trajectory of the dye attached to the hGBP1. Each dot is a dye state of a BD simulation lasting 8µs.

Peulen et al. (2017), J. Phys. Chem. B 121, 8211-8241

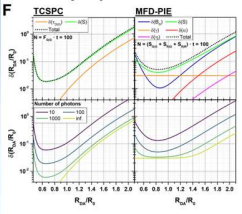
Brownian dynamics and Monte Carlo simulations



Peulen et al. (2017), J. Phys. Chem. B 121, 8211-8241

For average Donor-Acceptor distances the approximation error of the static dye model is less than 2 Å

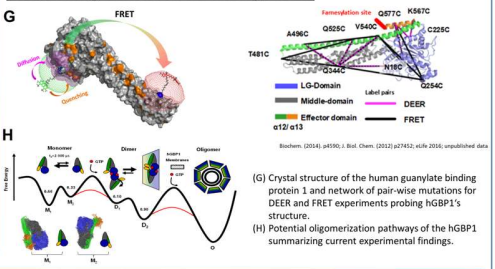
3. Error estimation in fluorescence lifetime and fluorescence intensity experiments



(F) The total uncertainty in the donor-acceptor distance assuming a normal distribution of the three main factors (1) the precision (noise) Δ_{noise}, (2) uncertainty of calibration Δ_{cal}, and (3) approximation error Δ_{model}.
α spectral fluorescence crosstalk, γ correction factor for quantum yields, η normalizing factor of donor/direct acceptor excitation rates, B_G background green channel, S total signal, $\tau_{D(0)}$ lifetime of donor only sample.

The error in fluorescence measurements only limits the measureable distances (giving an upper and lower limit)

Resolving structures using FRET



(G) Crystal structure of the human guanylate binding protein 1 and network of pair-wise mutations for DEER and FRET experiments probing hGBP1's structure.
(H) Potential oligomerization pathways of the hGBP1 summarizing current experimental findings.

References

[1] T. Peulen et al. (2017), J. Phys. Chem. B 121, 8211-8241
[2] M. Dimura et al. (2016), Curr. Opin. Struct. Biol., 40:163-185
[3] T. Vöpel et al (2014), Biochemistry, p. 4590
[4] E. Kravets et al. (2012) J. Biol. Chem., 27452-27466
[5] J. Lakowicz (2006), Principles of Fluorescence Spectroscopy, Springer

Acknowledgements

I would like to thank the International Helmholtz Research School of Biophysics and Soft Matter (IHRS BioSoft) and the European Research Council for financial support (ERC grant number 671208). And of course all members of my working group.

J. Folz, M.Popara, S. Felekyan and C.A.M. Seidel

"Determination of setup properties and calibration parameters using DNA"

Single-molecule Workshop, PicoQuant, Berlin 2019

Determination of setup properties and calibration parameters using DNA

J. Folz, M.Popara, S. Felekyan and C.A.M. Seidel
Heinrich-Heine-University, Düsseldorf

Heinrich Heine Universität Düsseldorf

BioSoft Biophysics and Soft Matter

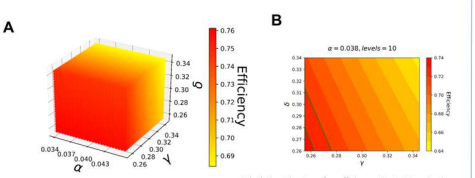
European Research Council

Established by the European Commission

Fluorescence spectroscopy and imaging are important biophysical techniques to study dynamics and function of biomolecules *in vitro* and in live cells. The use of more than one fluorophore per molecule opens additional opportunities arising from photon densities, coincidences and dipolar coupling by Förster Resonance Energy Transfer (FRET) to study the stoichiometry and structure of biomolecular systems.
Before performing FRET measurements, one needs to know the detection properties of the setup which depends on the photophysical properties of used dyes and characteristics of optical components. In the classical universal calibration via stoichiometry [1] these are mixed in one γ factor. Comparing multiple manifold γ factors leading to same FRET efficiencies raises the question which factor is the correct one. To answer this question and to improve accuracy we calculate the detection efficiencies using emission spectra of used dyes and spectra of every optical component in the setup. We used a rational determination to estimate the calibration factors using DNA samples and compared this to lifetime based analysis. This leads to a unique calibration that is valid for many molecules of interest with distinct conformers that are studied with the same donor/acceptor dye pair combination.

Problems/Uncertainties

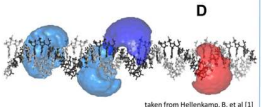
Calibration Parameter in intensity based FRET experiments



Calculating the transfer efficiency E via intensity is mainly based on three different calibration parameters, the crosstalk α of the donor emission in the red channels, the direct excitation δ of the donor laser to the acceptor dye, and γ, the ratio of detection efficiencies and quantum yields.
This results into a multiple set of calibration parameters resulting in the same transfer efficiency E. Figure **(A)** shows possible solutions for E in dependence of all three calibration parameters, **(B)** and **(C)** in dependence of only two, with one fixed.
It rises the question which set of parameter actually describes the physical situation. To testify and identify the calibration parameters individually we show an approach using DNA samples.

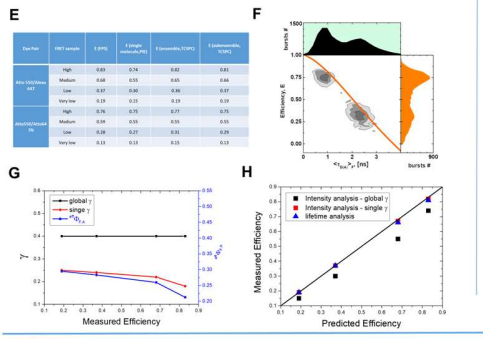
Global or single γ

Using a global γ is the usual and used approach in [1]. A global γ can only be calculated if you have more than one FRET species. However, for some dyes we experienced a high variation in the γ factor. As a model system we used the DNA **(D)** from [1] and tested several dyes and positions.

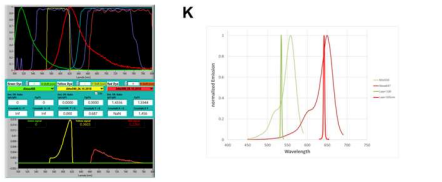


taken from Hellenkamp, B. et al [1]

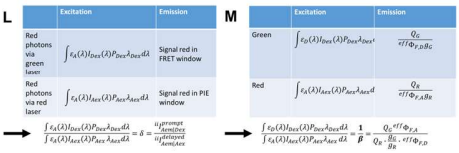
Using the dye Alexa 647 as an acceptor dye resulted in a discrepancy between intensity based FRET experiments and lifetime based FRET experiments. While the lifetime based approach matched in the transfer efficiency values with the predicted ones from FRET positioning screening (FPS) [2], the intensity approach failed. The measured efficiencies were too low with an increasing discrepancy with higher efficiencies (**E**, **F**).
We identified this shift in a change in the γ factor (**G**, **H**). To identify the reason for this shift a detailed calibration with certain parameters was needed.

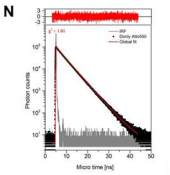


Solution – DNA calibration



(J) A dye (AttoRho13) emitting in both channels (green and red) has been used to calculate the ratio of detection efficiencies. The calculation is based on an overlay of the dye's spectrum and every optical component. Results have been compared to brightness measurements. Measuring the spectra of the donor and acceptor dye gives then the detection efficiency for the used dye pair
(K) Using the excitation spectra of the used dyes and the used lasers one can calculate the ratios of the excitation intensities **β**

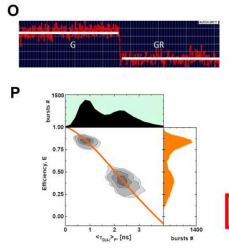




(L) Two methods to calculate the direct excitation δ are presented. One is over the excitation path, similar to the one shown in **(K)**. Another possible way is to use a acceptor only sample and compare the signals in the FRET (prompt) and PIE (delayed) window
(M) Two methods to calculate the ratios of excitation intensities **β**. One is over the excitation path, shown in **(K)**. Another possible way is over the emission path via the ratio of the brightnesses and quantum yields.
(N) Measuring the lifetime of the donor/acceptor attached to the protein (here is shown the fluorescence decay of Atto550) gives an estimate of the dye's quantum yield [3]. Due to environment effects like quenching the quantum yield could be reduced. Additionally the triplet states ratio has to be set into account (not shown here)

Using a few more simple measurements (acceptor only, donor only, a free dye emitting into all channels) results into a certain calibration of the setup

Resolving photophysical effects



Using single γ and calculated calibration parameters showed why the measured FRET efficiency in the intensity based approach differed from the ones in the lifetime approach: **(O)** shows the signal using a green laser (intervall marked with a G) and PIE with green and red laser (intervall marked with GR). A drop in the signal can be seen when the red laser is turned on under PIE conditions. The reason is a saturation effect due to the cis-trans states. This reduces the red quantum yield in dependance of the FRET efficiency (because the more energy transfer one has the more the acceptor dye saturates). After individually fixing the γ factor for the different FRET efficiency the values are on the static FRET line and match with the predictions from FPS.

Monitoring the photophysical behavior of the dye becomes in more detail is possible

References

[1] Hellenkamp, B., et al., *Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study.* Nature Methods, 2018. 15(11): p. 984-984.
[2] Kalinin, S., et al., *A toolkit and benchmark study for FRET-restrained high-precision structural modeling.* Nature Methods, 2012. 9(12): p. 1218-U129.
[3] Peulen, T.O., O. Opanasyuk, and C.A.M. Seidel, *Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately.* Journal of Physical Chemistry B, 2017. 121(35): p. 8211-8241.

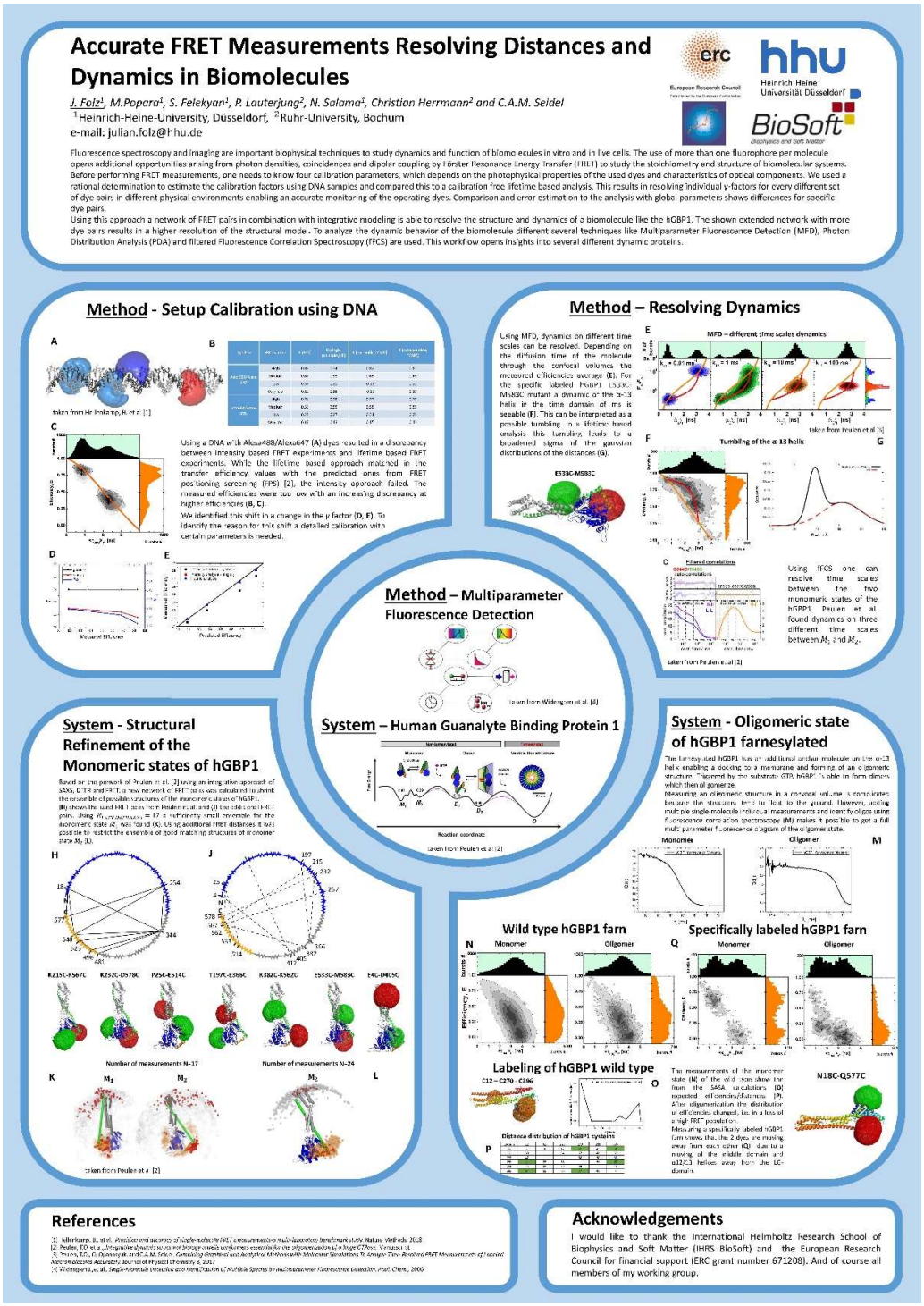
Acknowledgements

I would like to thank the International Helmholtz Research School of Biophysics and Soft Matter (IHRS BioSoft) and the European Research Council for financial support (ERC grant number 671208). And of course all members of my working group.

J. Folz, M.Popara, S. Felekyan, P. Lauterjung, N. Salama, C. Herrmann and C.A.M. Seidel

“Accurate FRET Measurements Resolving Distances and Dynamics in Biomolecules”

Biophysical Society Meeting, San Diego, 2020



6 Supplementary Information

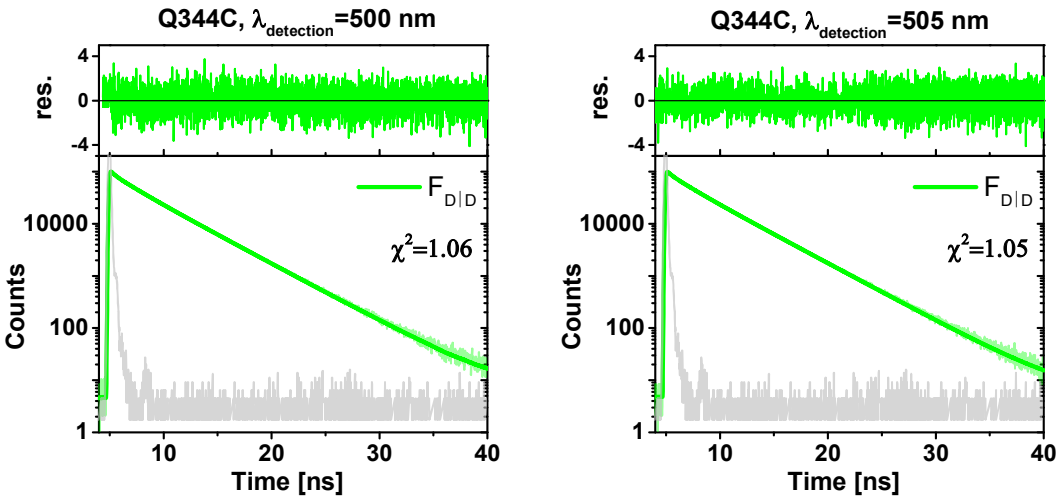
6.1 Using solvent relaxation to improve accuracy of smFRET measurements

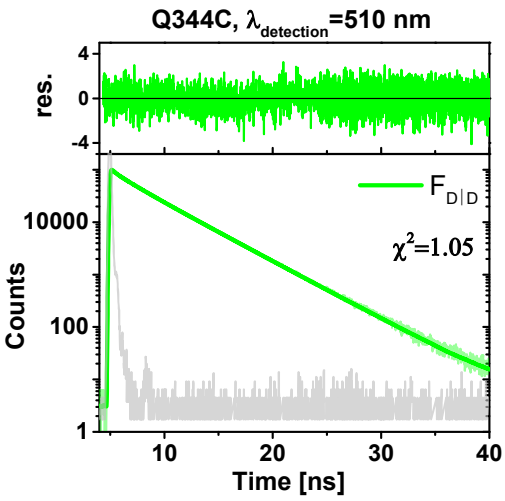
Methods

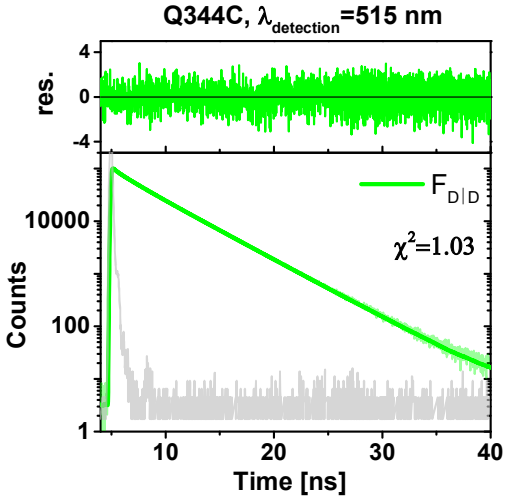
Data was taken on a FluoTime300 (PicoQuant, Berlin). Detection wavelength was automatically varied in 5 nm steps using a script. Sample temperature was held to 20 °C, measurements were taken to a maximum number of photons of 100000. For lifetime analysis measurements were taken under magic angle for detection, for anisotropy measurements were taken under 0° and 90° angles. Lifetime analysis was done using ChiSurf available under https://github.com/Fluorescence-Tools/chisurf and described in literature [44]. For anisotropy analysis Matlab was used using scripts available under *P:\Solvent_relaxation_jf.*

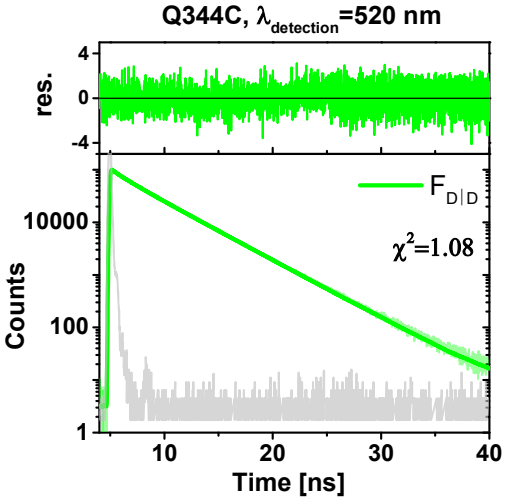
Data

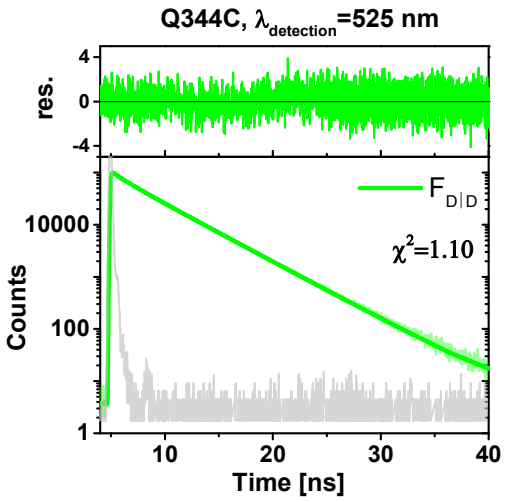
hGBP1 Q344C-Alexa488

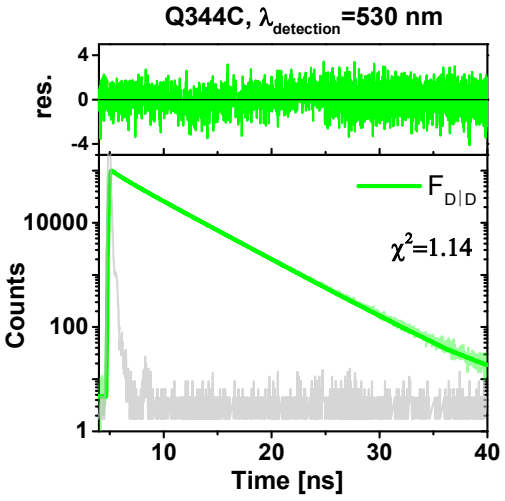


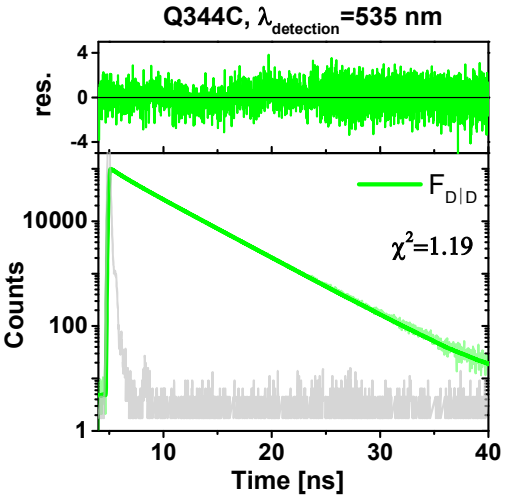

Q344C, λdetection=510 nm
res.
FD|D
χ2=1.05
Counts
Time [ns]

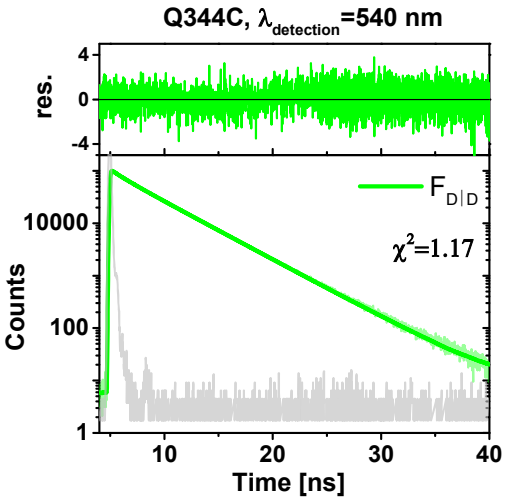

Q344C, λdetection=515 nm
res.
FD|D
χ2=1.03
Counts
Time [ns]

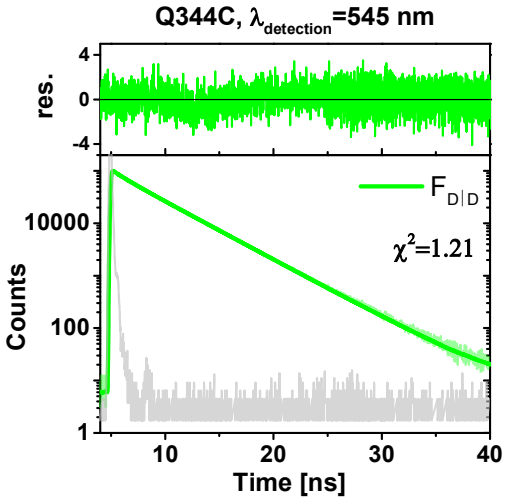

Q344C, λdetection=520 nm
res.
FD|D
χ2=1.08
Counts
Time [ns]

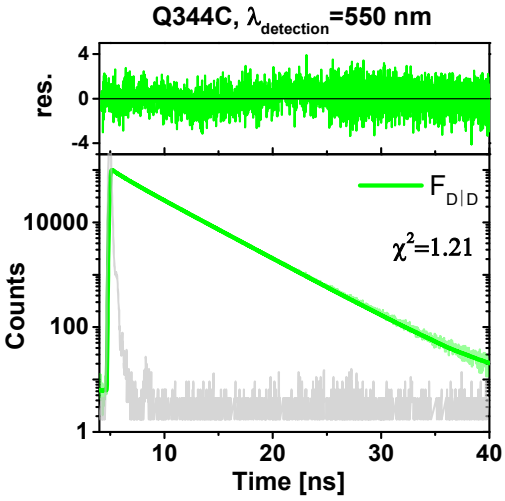

Q344C, λdetection=525 nm
res.
FD|D
χ2=1.10
Counts
Time [ns]

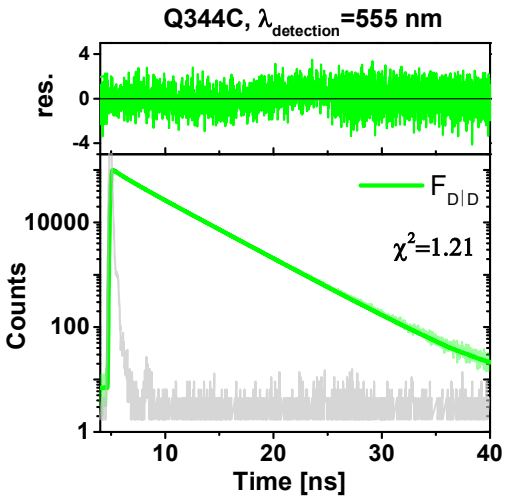

Q344C, λdetection=530 nm
res.
FD|D
χ2=1.14
Counts
Time [ns]

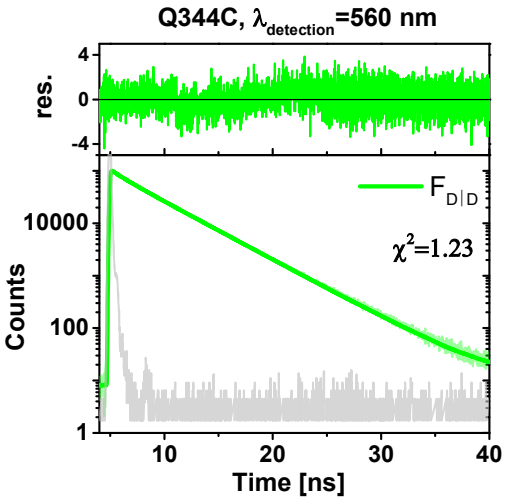

Q344C, λdetection=535 nm
res.
FD|D
χ2=1.19
Counts
Time [ns]

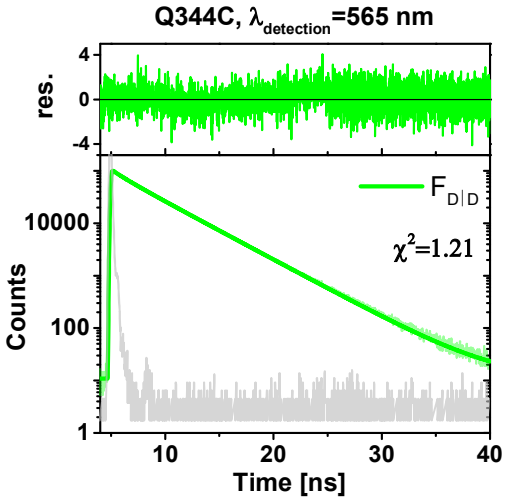

Q344C, λdetection=540 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.17
10
20
30
40
Time [ns]

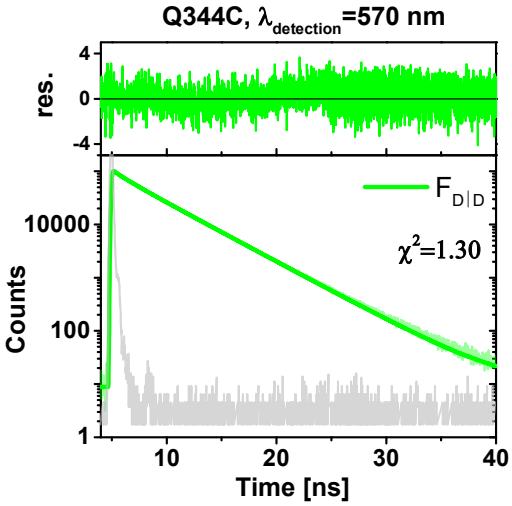

Q344C, λdetection=545 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.21
10
20
30
40
Time [ns]

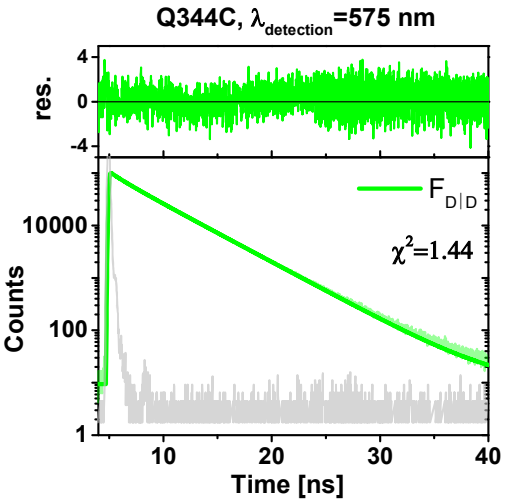

Q344C, λdetection=550 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.21
10
20
30
40
Time [ns]

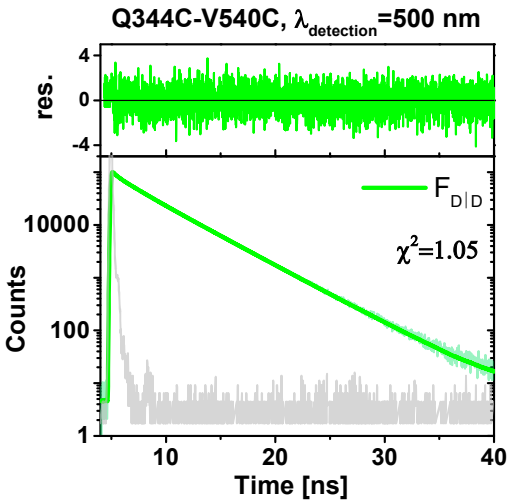

Q344C, λdetection=555 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.21
10
20
30
40
Time [ns]

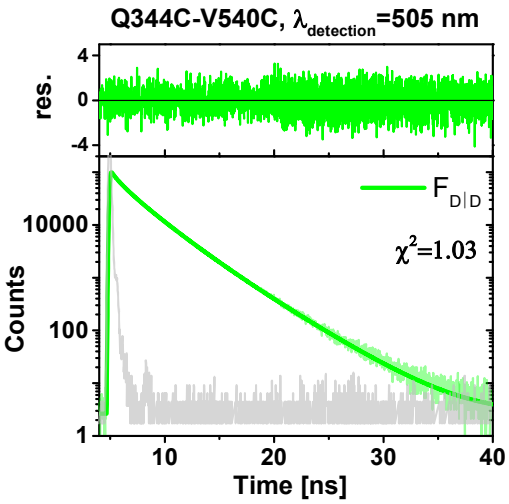

Q344C, λdetection=560 nm
res.
4
0
-4
Counts
10000
100
1
10
20
30
40
Time [ns]
FD|D
χ2=1.23

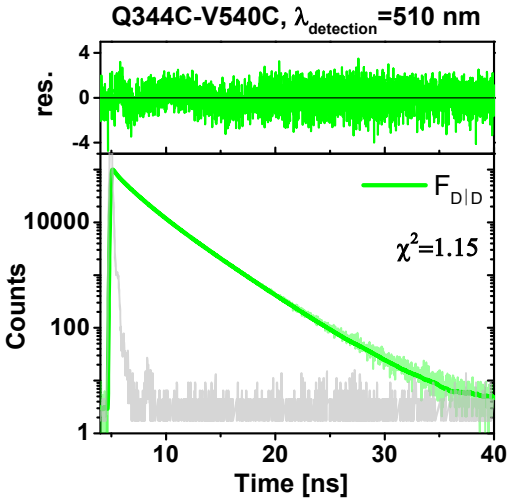

Q344C, λdetection=565 nm
res.
4
0
-4
Counts
10000
100
1
10
20
30
40
Time [ns]
FD|D
χ2=1.21

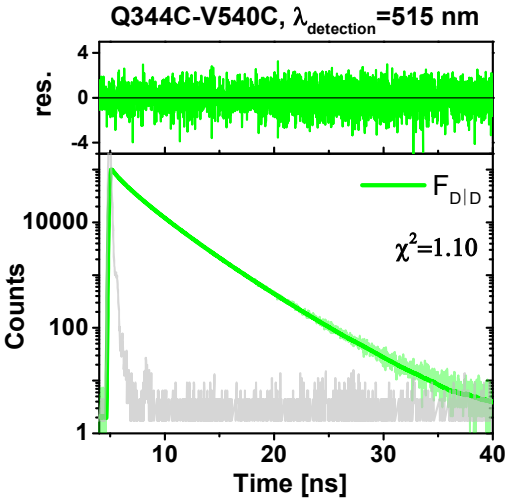

Q344C, λdetection=570 nm
res.
4
0
-4
Counts
10000
100
1
10
20
30
40
Time [ns]
FD|D
χ2=1.30

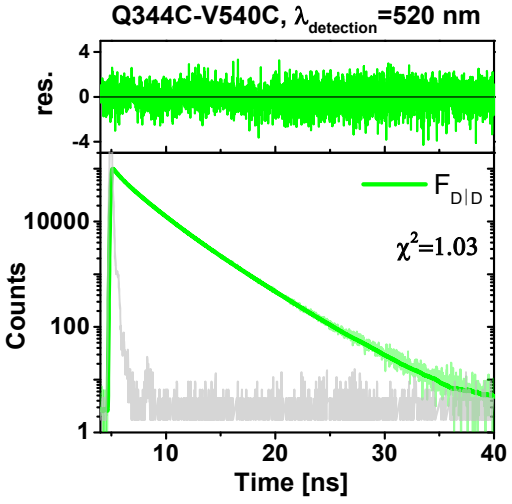

Q344C, λdetection=575 nm
res.
4
0
-4
Counts
10000
100
1
10
20
30
40
Time [ns]
FD|D
χ2=1.44

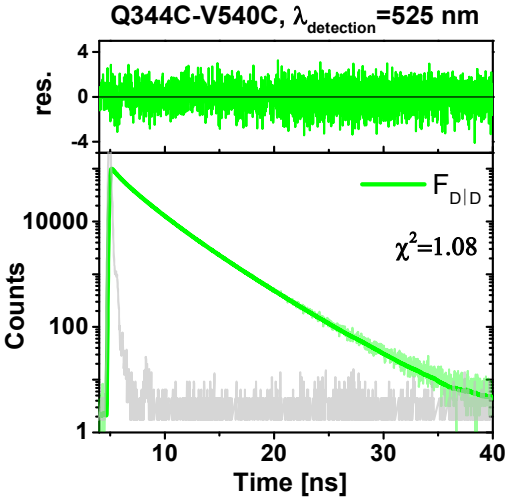

Q344C-V540C, λdetection=500 nm
res.
Counts
Time [ns]
FD|D
χ2=1.05

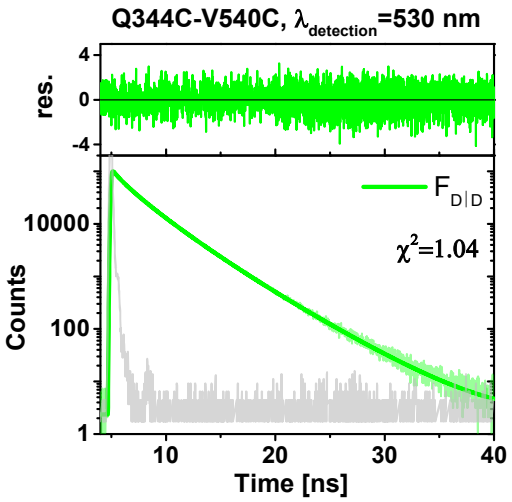

Q344C-V540C, λdetection=505 nm
res.
Counts
Time [ns]
FD|D
χ2=1.03

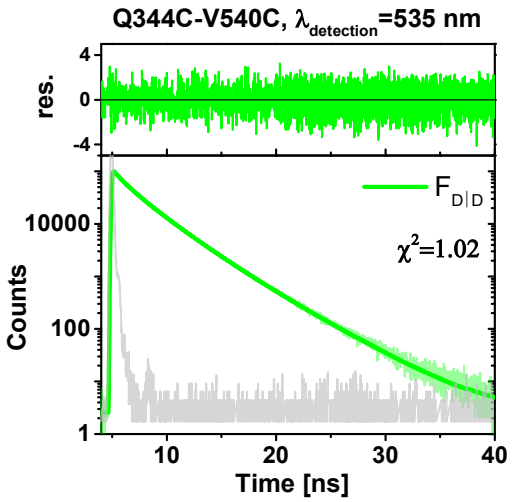

Q344C-V540C, λdetection=510 nm
res.
Counts
Time [ns]
FD|D
χ2=1.15

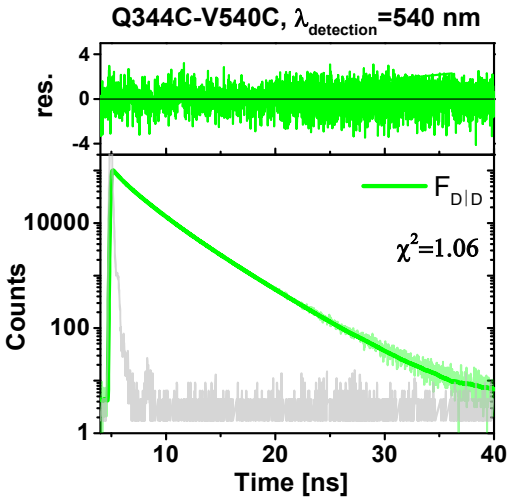

Q344C-V540C, λdetection=515 nm
res.
Counts
Time [ns]
FD|D
χ2=1.10

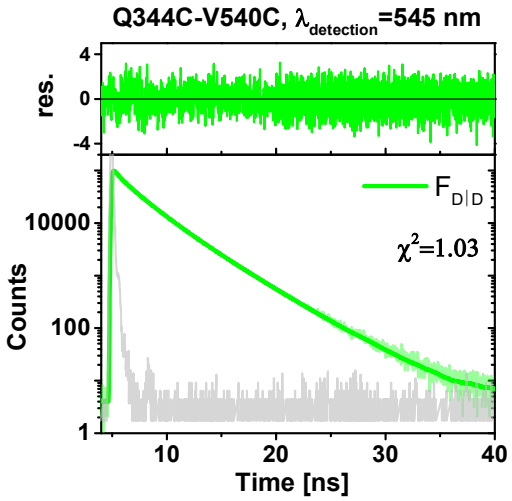

Q344C-V540C, λdetection=520 nm
res.
Counts
Time [ns]
FD|D
χ2=1.03

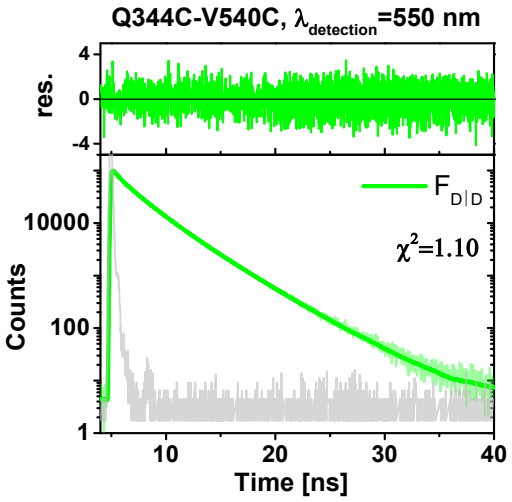

Q344C-V540C, λdetection=525 nm
res.
Counts
Time [ns]
FD|D
χ2=1.08

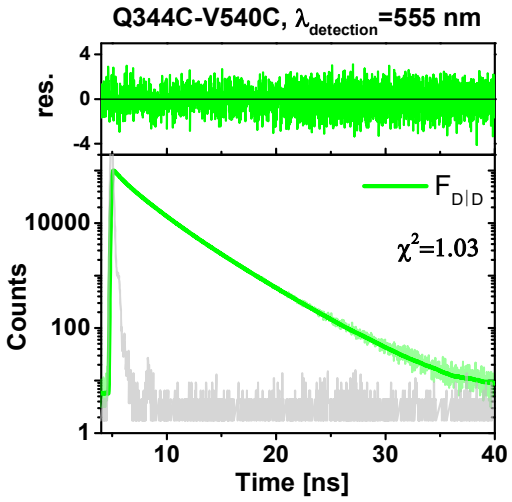

Q344C-V540C, λdetection=530 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.04
10
20
30
40
Time [ns]

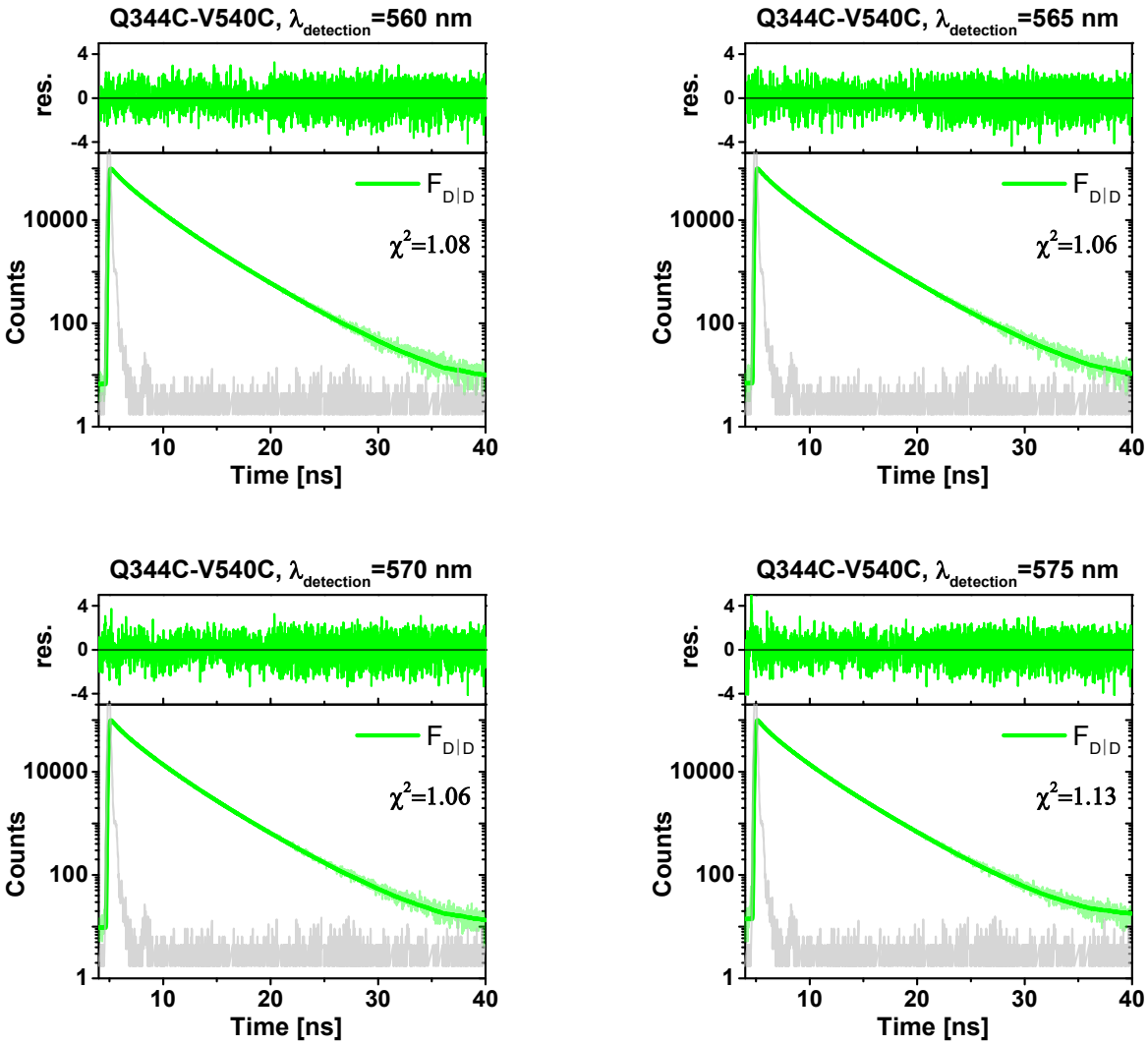

Q344C-V540C, λdetection=535 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.02
10
20
30
40
Time [ns]


Q344C-V540C, λdetection=540 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.06
10
20
30
40
Time [ns]


Q344C-V540C, λdetection=545 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.03
10
20
30
40
Time [ns]


Q344C-V540C, λdetection=550 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.10
10
20
30
40
Time [ns]


Q344C-V540C, λdetection=555 nm
res.
4
0
-4
Counts
10000
100
1
FD|D
χ2=1.03
10
20
30
40
Time [ns]



6.2 Resolving the mechanism behind pore formation of the Tc toxin cia an analysis of states, rates and transition times

Methods

single-molecule confocal

Single-molecule measurements with Multiparameter Fluorescence Detection (MFD) were performed on a home built setup based on an Olympus IX70 inverted microscope as described in [106]. An Olympus UPlanSAp 60x/1.2 objective was used. A linearly polarized, pulsed diode laser with a wavelength of 495 nm (LDH-D-C 495, PicoQuant) operated at 64 MHz was used. In case of PIE configuration () an additional red laser with a wavelength of 635 nm was used, both operated at 32 MHz. For detection, the beam is splitted into parallel and perpendicular polarization and filtered by color using ET 535/50 and HQ 730/140 (AHF Analysentechnik) and finally detected by 8 detectors. Single photon counting was done with synchronized channels (HydraHarp 400, PicoQuant, Germany) operating in Time-Tagged Time-Resolved (TTTR) mode. Burst selection

and data analysis was done using established procedures and in house software described in [107], available upon request on the homepage of the Seidel group (https://www.mpc.hhu.de/software.html). To prevent concentration drops due to sticking of the molecule the cover glass was passivated with BSA using 10µM stock solution which was washed off after a 2 minutes incubation time.

Single-molecule events were identified using a burst search algorithm according to [40] using a Lee filter, a threshold of 0.2 ms and a minimum of 60 photons per burst. For anisotropy PDA the whole trace including all bursts were analyzed. For heteroFRET analysis, molecules with a stoichiometry between *S*=0.5 to *S*=0.8 were selected. The labeling geometry of 5 possible labeling spots for TcA(914-BFL)-Biotin:TcB(1041-Atto647N)-TcC and labeling scheme of 5:1 donor to acceptor lead to a higher than usual value of the stoichiometry, where multiple donors where attached to the molecule with only one acceptor. However, only one of the label position have a distance to the acceptor where FRET occurs, giving the possibility to qualitatively measure FRET efficiency derived distances.

Static FRET lines as described in [32] were used as a consistency and quality control. A static FRET line relates the lifetime of the donor in presence of an acceptor to the resulting FRET efficiency. In case of the TcToxin no dynamic exchange between conformational states is expected. This matches the experimental data, where the double labeled FRET population of the TcToxin in PrePore and Pore state on the $\langle \tau_{D(A)} \rangle_F - E$ is located on the static FRET line.

single-molecule TIRF

Imaging was performed on a home built Total Internal Reflection Fluorescence (TIRF) setup based on Olympus IX70 microscope and CCD camera (emCCD, DV887 ECS-BV, Andor). Using a special TIRF objective (Apo N 60x, 1.49 NA, Olympus) with very high numerical aperture leads to total internal reflection of the light beam on the surface of the glass. For homoFRET-studies a single color excitation with a 488 nm continuous wave laser (Cobolt MLD) was used. In case of hetero FRET studies an additional 635 nm cw-laser (Cobolt MLD) was added using alternating excitation (ALEX), where excitation was alternating triggered with a home-built switch. Emission light passes a dichroic mirror and is then splitted by polarization (homoFRET) or by colors (heteroFRET) and projected as two spatially separated images (parallel/perpendicular or green/red, respectively) on the camera using an image splitter (OptoSplit II, Cairn Research Ltd). Splitting by colors was done using (HC BS 580 Imaging, AHF Analysentechnik) and bandpass filters (green: HQ 535/50, red: HQ 680/60), splitting by polarization via polarizing beamsplitting cube. For every measurements, 2048 images with a pixel size of 512x512 were taken with a single frame time of 29.44 ms. Spot selection and data analysis was done using the software iSMS [108]. For background subtraction mean intensity of a background mask was used. Spot selection was done using an automatized intensity based algorithm. To calibrate the overlay of the two individual images a measurement with inhomogeneous areas was used. Polarization traces are smoothed using Savitzky-Golay algorithm with varying number of points. Efficiency traces are averaged based on the time interval before and after the transition.

Total internal reflection fluorescence (TIRF) uses an evanescent field coming from the high refractive index difference of the TIRF objective and the sample. This results in a total reflection of an incoming laser beam in case it enters the medium with an angle higher than a critical angle based on the difference of the refractive indices following Snell's law. Thus, it makes the technique ideal for samples immobilized to the surface, because the laser beam only enters typically around 100 nm of the solution, leading to a high signal to noise ratio.

The correction factors were calculated based on the spectral properties of the dyes and applied globally for every molecule. Following correction factors were used:

Supplementary Table 1: Overview of calibration parameters used for homo and hetero FRET

	HeteroFRET Alexa488/Atto647N	homoFRET Atto647N/Atto647N
α	0.016	0
δ	0.02	0
$\frac{g_{R\|A}}{g_{G\|D}}$	1.25	1
$\frac{^{eff}\Phi_{F,A}}{^{eff}\Phi_{F,D}}$	$\frac{0.65}{0.8} = 0.81$	1
γ	1.02	1

For heteroFRET-efficiency background corrected fluorescence *F* was used following equations described in section intensity based MFD.

In case of the homoFRET anisotropy essay one has to take into account that the high numerical aperture of the TIRF objective influences the polarization of the laser beam, leading to a mixture of polarizations. Additionally the laser beam enters the solution at an angle greater than a critical angle for total reflection. This effect was studied in detail and a shift to lower anisotropy values was observed (see e.g. [109], [110]). Additionally, due to a high amount of different labeling sites of the biotin on the TcToxin, the initial orientation and therefore polarization values of the TcToxin are distributed. However, in this essay the difference of polarization of the signal before and after the transition of the TcToxin from Prepore to Pore was of interest. Therefore, we applied a polarization offset based on the mean polarization value before the pH change $\langle P(\mathrm{pH7})\rangle$ to every trace shifting the polarization value before the pH change to *P* = 0. The polarization value results in:

$$P = \frac{F_{||} - F_{\perp}}{F_{||} + F_{\perp}} - \langle P(\mathrm{pH7})\rangle \qquad (6.2\text{-}1)$$

Immobilization

TcToxins were immobilized on the surface using biotinylated BSA. Cleaning of nunc chambers (Nunc Lab-Tek II, thickness No. 1.5H, ThermoFisher) was done by activating them in an oxygen plasma for 2 min (FEMTO Plasma Cleaner, Diener electronic). After cleaning the surface was incubated for 10 min with biotinylated BSA (Sigma-Aldrich, 3 mg/ml in PBS), washed several times with PBS, incubated for 10 min with neutravidin (Invitrogen, 20µg/ml in water), and washed again with buffer. The biotinylated Toxins were added and incubated for up to 30 min. Finally, the chamber was washed again to remove diffusing toxins and dyes.

Correction factors of intensity based confocal MFD

Correction factors were estimated from the spectra of the used dyes obtained from the web/manufacturer [111] and the spectra of the optical components based upon [112]. The spectral donor crosstalk of the donor α is defined as the detection efficiency ratios *g* of the red to green detection channel while donor excitation:

$$\alpha = \frac{g_{R|D}}{g_{G|D}} \tag{6.2-2}$$

The direct excitation of the acceptor with the donor laser δ is based on the ratio of the cross-sections of the acceptor excited with the donor laser $\sigma_{A|G}$ and with the acceptor laser $\sigma_{A|R}$ and the laser intensities:

$$\delta = \frac{\sigma_{A|G}}{\sigma_{A|R}} \frac{I_{Dex}}{I_{Aex}} \tag{6.2-3}$$

To monitor the labeling ratio of donor to acceptor of the molecule the excitation flux are normalized using

$$\beta = \frac{\sigma_{A|R}}{\sigma_{D|G}} \frac{I_{Aex}}{I_{Dex}} \tag{6.2-4}$$

what from follows that molecules labeled with a 1:1 ratio have a stoichiometry value *S* = 0.5. Due to the complex labeling geometry of possible multilabeling an additional ALEX-2CDE filter described in [41] was applied to remove further unwanted contributions of single labeled molecules and reduce the effect of photobleaching. Differences in the quantum yields and the detection efficiencies of the different spectral ranges lead to γ correction factor defined as:

$$\gamma = \frac{g_{R|A}}{g_{G|D}} \frac{^{eff}\Phi_{F,A}}{^{eff}\Phi_{F,D}} \tag{6.2-5}$$

with the dark state corrected effective quantum yields of the acceptor and donor.

Intensity based spectroscopic parameters

Raw intensities for green and red signal after donor and acceptor excitation are corrected from background taken from a only buffer measurement:

$$ {}^{ii}I_{\mathrm{Dem|Dex}} = {}^{i}I_{\mathrm{Dem|Dex}} - I^{(BG)}_{\mathrm{Dem|Dex}} \tag{6.2-6} $$

$$ {}^{ii}I_{\mathrm{Aem|Dex}} = {}^{i}I_{\mathrm{Aem|Dex}} - I^{(BG)}_{\mathrm{Aem|Dex}} \tag{6.2-7} $$

$$ {}^{ii}I_{\mathrm{Aem|Aex}} = {}^{i}I_{\mathrm{Aem|Aex}} - I^{(BG)}_{\mathrm{Aem|Aex}} \tag{6.2-8} $$

and furthermore corrected from the calculated correction parameters to the fluorescence signal

$$ F_{\mathrm{D|D}} = \gamma\, {}^{ii}I_{\mathrm{Dem|Dex}} \tag{6.2-9} $$

$$ F_{\mathrm{A|A}} = \frac{1}{\beta}\, {}^{ii}I_{\mathrm{Aem|Aex}} \tag{6.2-10} $$

$$ F_{\mathrm{A|D}} = {}^{ii}I_{\mathrm{Aem|Dex}} - \alpha\, {}^{ii}I_{\mathrm{Dem|Dex}} - \delta\, {}^{ii}I_{\mathrm{Aem|Aex}} \tag{6.2-11} $$

which can now be used to calculate FRET efficiency *E* and the stoichiometry *S*:

$$ E = \frac{F_{\mathrm{A|D}}}{F_{\mathrm{D|D}} + F_{\mathrm{A|D}}} \tag{6.2-12} $$

$$ S = \frac{F_{\mathrm{D|D}} + F_{\mathrm{A|D}}}{F_{\mathrm{D|D}} + F_{\mathrm{A|D}} + F_{\mathrm{A|A}}} \tag{6.2-13} $$

Förster Radius was calculated using the spectral overlap of the emission of the donor and the excitation of the acceptor dye J. For the dipole orientation factor κ^2 the isotropic average of 2/3 was assumed, for the refractive index a commonly used value of *n*=1.4 was taken. Following this the FRET efficiency averaged distance [113] $\langle R_{DA}\rangle_E$:

$$ \langle R_{DA}\rangle_E = R_0 (E^{-1} - 1)^{1/6} \tag{6.2-14} $$

FRET Position Screening (FPS)

In order to design a label scheme with dyes positioned in the FRET sensitive range to measure the conformational change due to the shell changing and the pore injection, a FRET position screening was applied following [17] using in house software. In short, it applies positionwise so called Accessible Volumes (AVs) [48] to multiple positions on the molecule while calculating the expected distance $\langle R_{DA}\rangle_E$. To do so, the first position for the donor dye was set to an already used position for other experiments to reduce time and effort. However, after suggestion of different label scheme specific selections were done considering the molecule's environment of the position, then labelled and finally tested using smMFD. The outcome was that the most sufficient label scheme in terms of good label ratio for the shell changing was the TcA(1279-Atto647N) and for the pore injection the TcA(914-BFL:TcB(1041-Atto647N)-TcC.

Anisotropy Photon Distribution Analysis

Anisotropy states and fractions were estimated via PDA. Anisotropy is changed based on different phenomena. In case of TcA-2365-Bdp and TcA-2384-BdP an environmental change occurring from the transition from PrePore to Pore state is restricting the dye's movement in Pore state, leading to a higher anisotropy. In case of TcA-1193-Bdp and TcA-1279-Atto647N the anisotropy in Pore state is decreased due to a higher distance of the dyes and therefore the reduction of homoFRET Efficiency. Based on [114], [107] anisotropy PDA is performed. There, the sum of the probability of every combination of the ratio $S_{||}/S_{\perp}$ is

$$P\left(\frac{S_{||}}{S_{\perp}}\right) = \sum_{\left(S_{||}/S_{\perp}\right)_i} P(S_{||}, S_{\perp}, B_{||}, B_{\perp}) \tag{6.2-15}$$

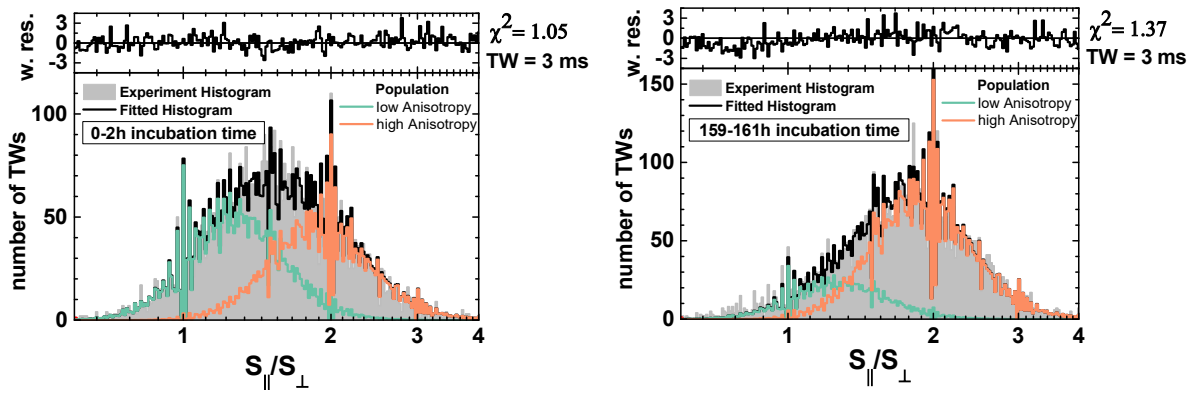
where $S_{||}$ is the parallel polarized signal and $S_{\perp}$ the perpendicular polarized signal, B their corresponding background signal. Experimental data is following Poisson statistics and fitted using a maximum entropy method described in the given references. The scatter-corrected fluorescence follows:

$$r_s = \frac{G\left(S_{||} - \langle B_{||}\rangle\right) - \left(S_{\perp} - \langle B_{\perp}\rangle\right)}{G\left(S_{||} - \langle B_{||}\rangle\right)(1 - 3l_2) + \left(S_{\perp} - \langle B_{\perp}\rangle\right)(2 - 3l_1)} \tag{6.2-16}$$

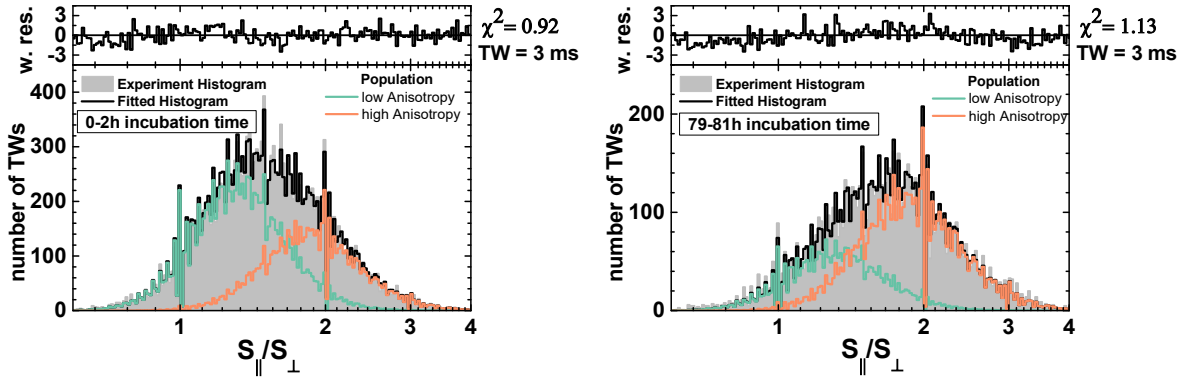
with the correction factors l_1 and l_2 describing the mixing of the polarizations in the microscope objective and G the ratio of detection in parallel and perpendicular channel, obtained from a free dye measurement. PDA analysis was performed globally over 3 different time steps, 1 ms, 2 ms and 3 ms. The experimentally number of time windows with a particular r_s were fitted using a 2-state model with a high and low anisotropy state. Fitting involved minimization of reduced χ^2-values. The resulting time dependent fraction of high and low anisotropy states are then describing the kinetic behavior of the TcToxin.

Supplementary figures:

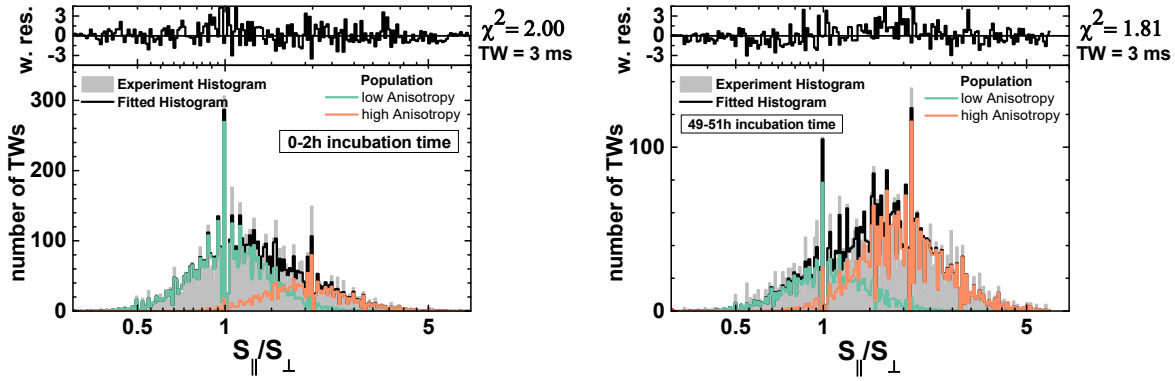
a – TcA(2365-BFL)



b – TcA(2384-BFL)



c – TcA(1193-BFL)

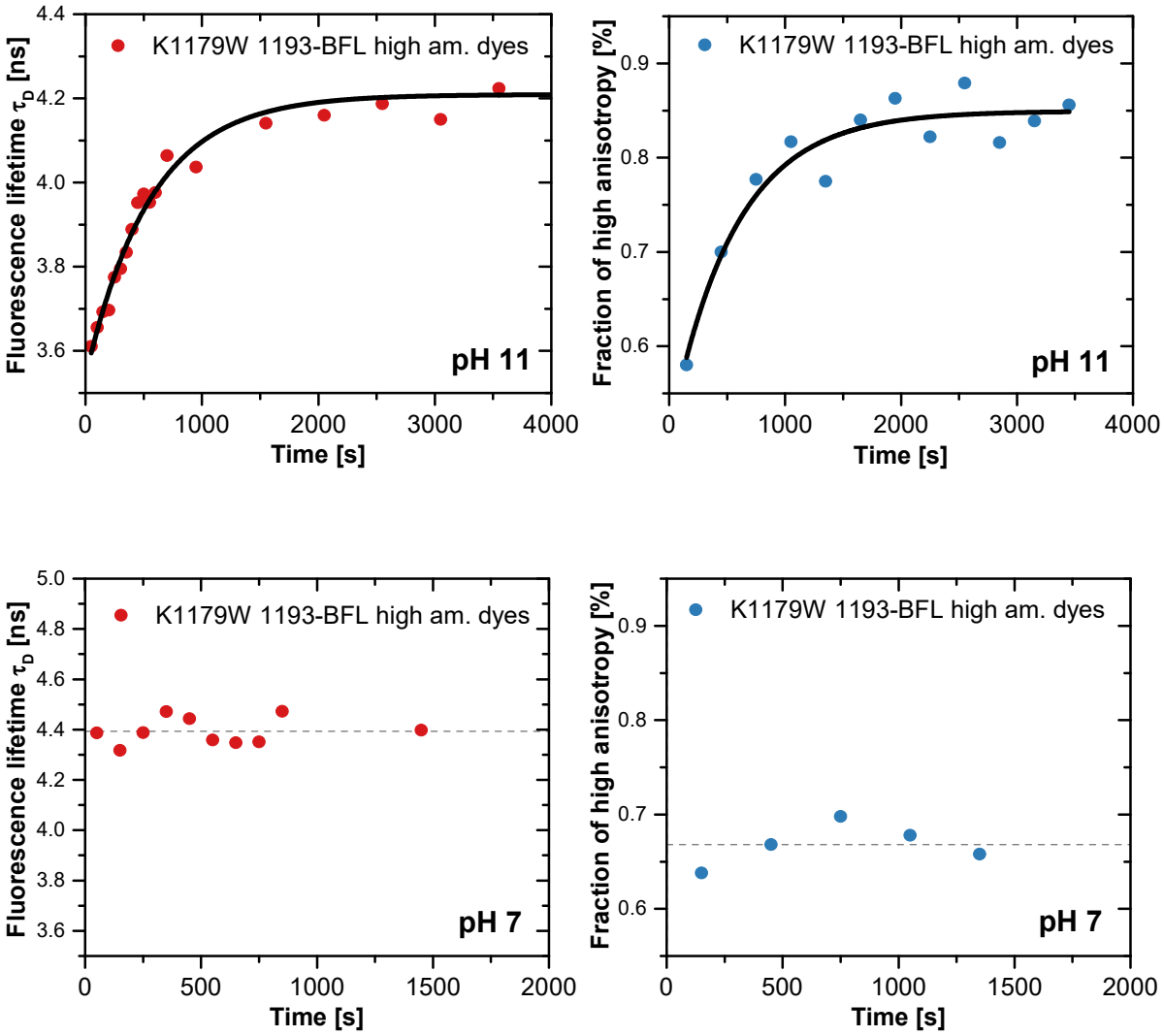


Supplementary Figure 1: Anisotropy Photon Distribution Analysis with a time window TW = 3 ms of freely diffusing single TcA subunits. Start- and endpoint of long time measurement series shown in main figure 1 of a TcA(2365-BFL) and b TcA(2384-BFL) monitoring injection, and c TcA(1193A-BFL) monitoring shell opening. The data are well described by a 2 state model (black) with a superposition of a low anisotropy state (lr, green) and high anisotropy state (hr, orange). The quality of the fit was judged by weighted residuals (upper panels) and X2. The fit results are compiled in Table SI 1. In case of TcA(2365-BFL) and TcA(2384-BFL) high anisotropy state (hr = 0.29) is caused by environmental change from the surrounding opened

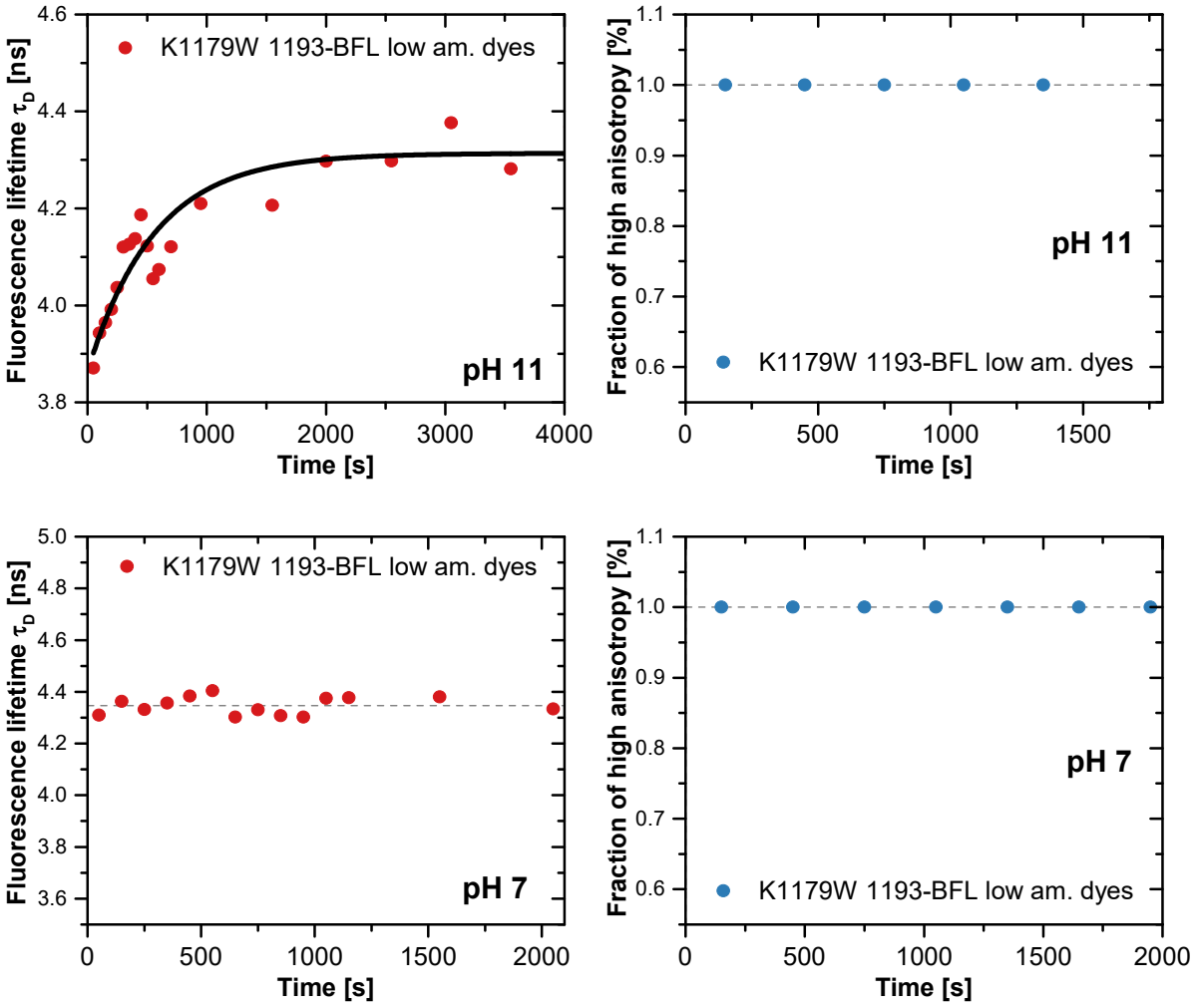
shell/injected pore therefore interpreted as Pore state. For TcA(1193-BFL) high anisotropy state (r = 0.27) is caused by environment change and reduction of homoFRET efficiency, indicating a destabilization of the shell

Supplementary Table 2: Photon Distribution Analysis. Fit-Results of the displayed measurements/analysis.

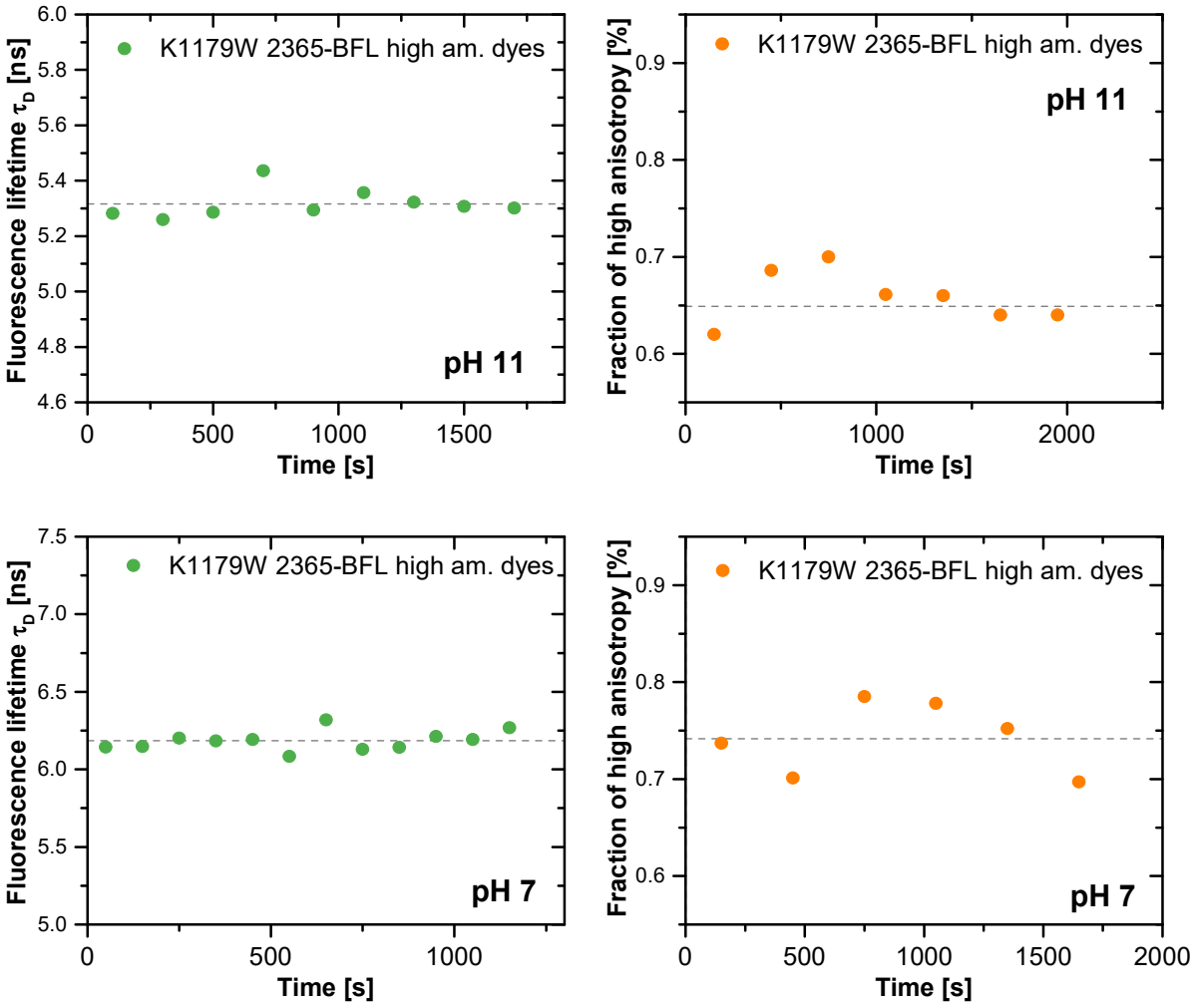
	High anisotropy	Low anisotropy	Incubation time	Fraction of high anisotropy state [%]	Fraction of low anisotropy state [%]
TcA(2365-BFL)	0.29	0.13	1	48.8	51.2
			160	78.1	21.9
TcA(2384-BFL)	0.29	0.14	1	39.3	60.7
			80	66.8	33.2
TcA(1193-BFL)	0.27	0.03	1	29.2	70.8
			50	67.5	32.5



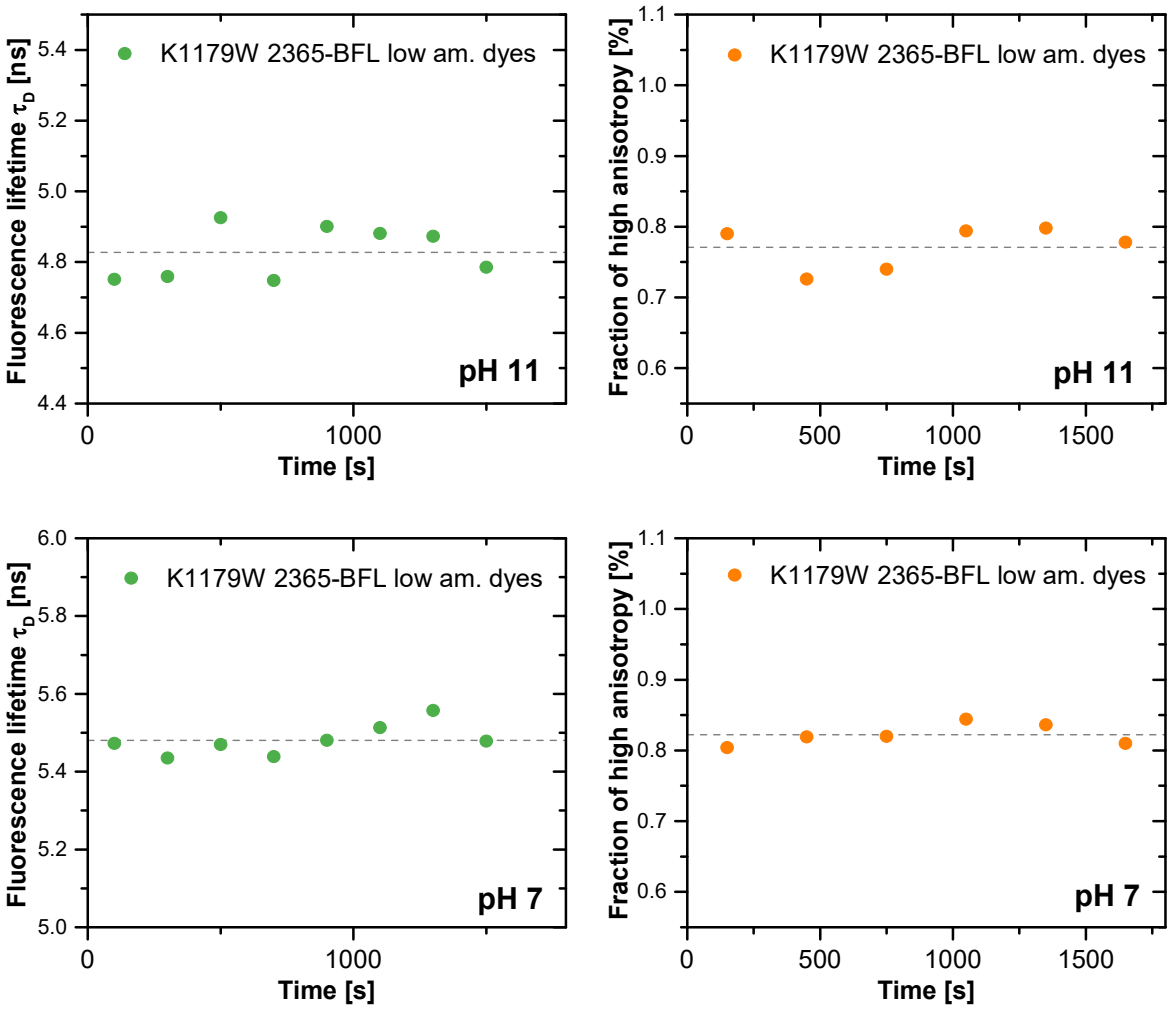
Supplementary Figure 2: Measurements of the intermediate mutant K1197W labeled with a high amount of dyes. The fluorescence lifetime in dependence of the time was fitted using $\boldsymbol{\tau_D(t) = Ae^{-kt} + y}$ with *A* the amplitude, *y* the offset and *k* the rate of the kinetic. As a result, a rate of 0.00179 was fitted for both kinetics globally. This corresponds to a relaxation time of 558 seconds.



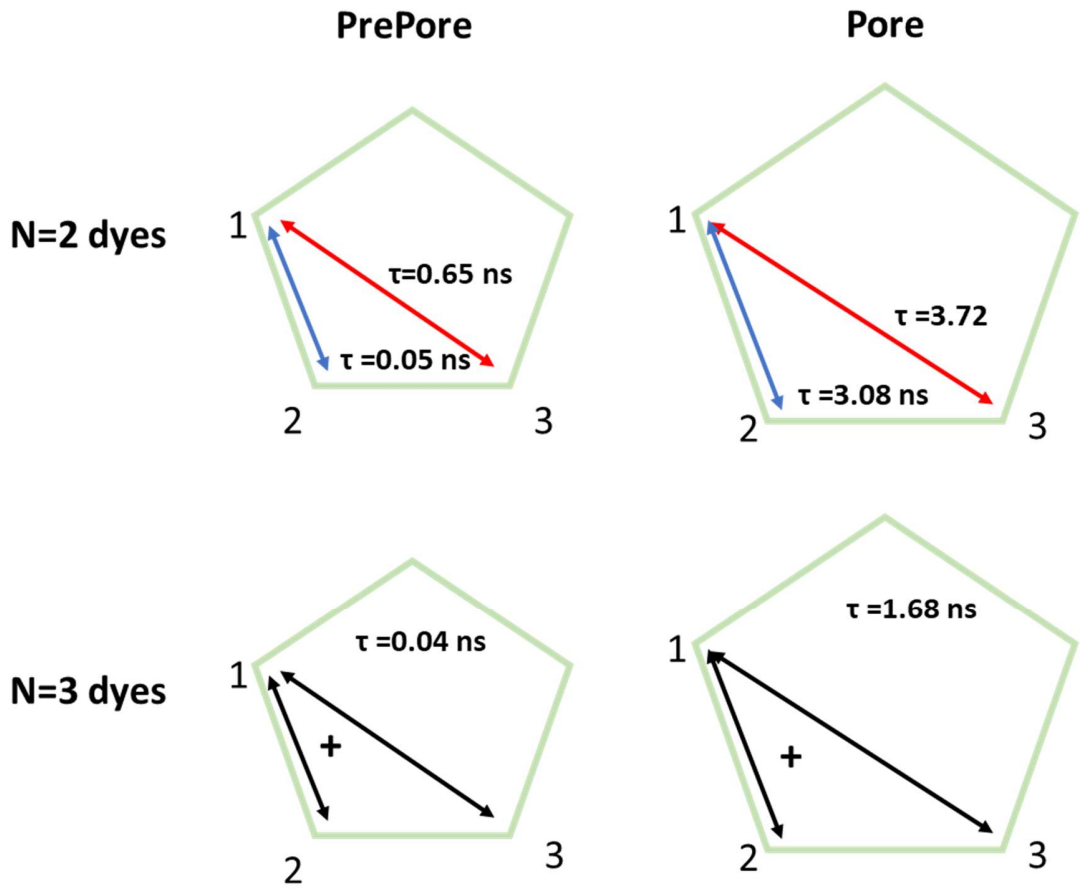
Supplementary Figure 3: Measurements of the intermediate mutant K1197W labeled with a low amount of dyes. The fluorescence lifetime in dependence of the time was fitted using $\boldsymbol{\tau_D(t) = Ae^{-kt} + y}$ with *A* the amplitude, *y* the offset and *k* the rate of the kinetic. As a result, a rate of 0.00179 was fitted for globally with the other kinetic shown in **Supplementary Figure 3**. This corresponds to a relaxation time of 558 seconds.



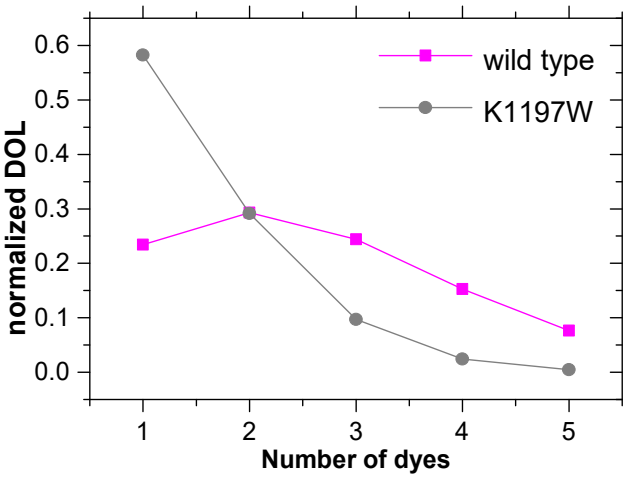
Supplementary Figure 4: Measurements of the intermediate mutant K1197W labeled with a high amount of dyes.



Supplementary Figure 5: Measurements of the intermediate mutant K1197W labeled with a low amount of dyes.



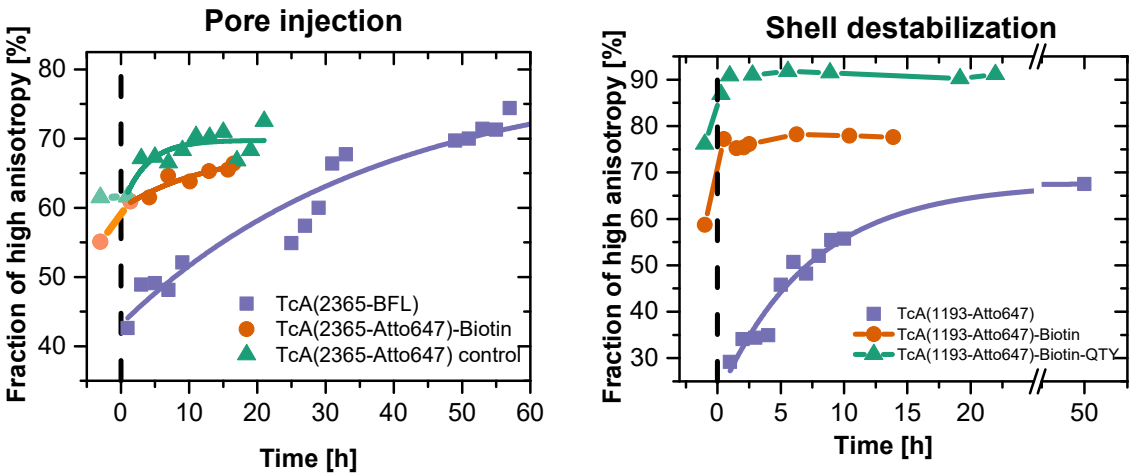
Supplementary Figure 6: Labeling scenarios for different degree of labeling. For two dyes, the label can be either the next neigbhour, or diagnolly orientated. For three dyes, they are always labeled to the next neigbour and diagonally orientated.



Supplementary Figure 7: Poissonian distribution of the number of labeled dyes. For the wild type a high degree of labeling was measured (average number of dyes is 2.5), whereas for the K1197W mutant only an average of 1 dye per molecule was estimated.

Supplementary Table 3

Settings		Experiment			Expectation		
Mutant	**Incubation at pH 11**	**Ampl. fast Depol. [%]**	**Ampl. slow Depol. [%]**	**FRETactive molecules [%]**	**Ampl. fast Depol. [%]**	**Ampl. slow Depol. [%]**	**FRETactive molecules (≥2 dyes) [%]**
wild type	**start**	100	0	71	100	0	77
	end	50.5	49.5	50	62	38	77
intermediate	**start**	66	34	42	100	0	42
	end	62	38	32	30	70	42

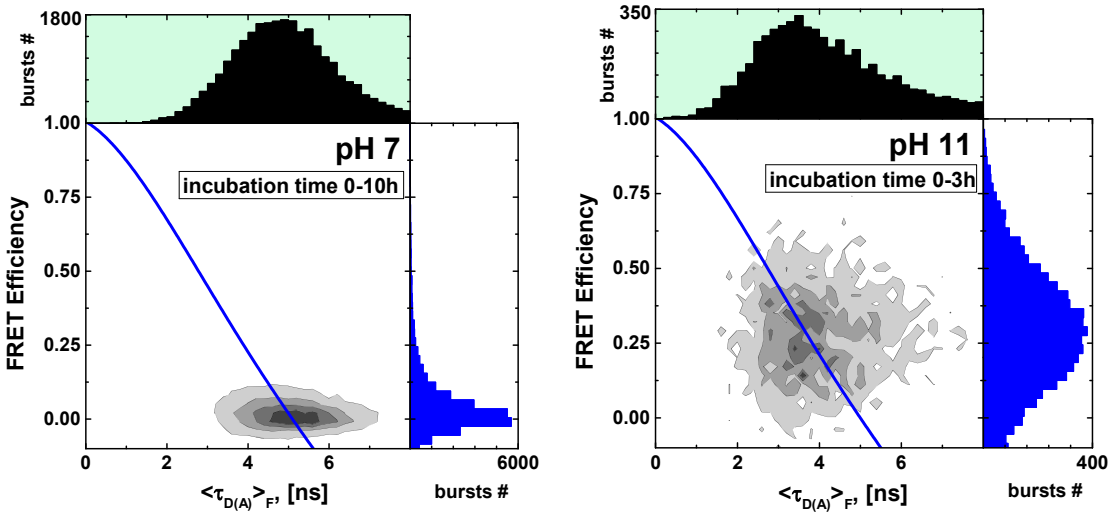


Supplementary Figure 8: Biotin kinetics. Confocal single-molecule measurements using PDA analysis show different kinetic behavior for biotinylated toxins. On the left hand side the pore injection is monitored with the TcA(2365-BFL) sample in comparison to a biotinylated TcA(2365-Atto647N) and a non-biotinylated TcA(2365-Atto647N) sample. It is seeable that using the biotinylated sample most of the kinetics/change of fraction from PrePore to Pore state happens shortly after the pH-change (dashed black line). On the right hand side the shell destabilization is monitored and is showing the same effect. Where the non-biotinylated sample shows a large fraction change, the biotinylated TcA(1193-Atto647N) sample shows an instant (less than several minutes) fraction change after pH change. The same behavior was observed at the TcA(1193-Atto647N)- Biotin QTY sample.

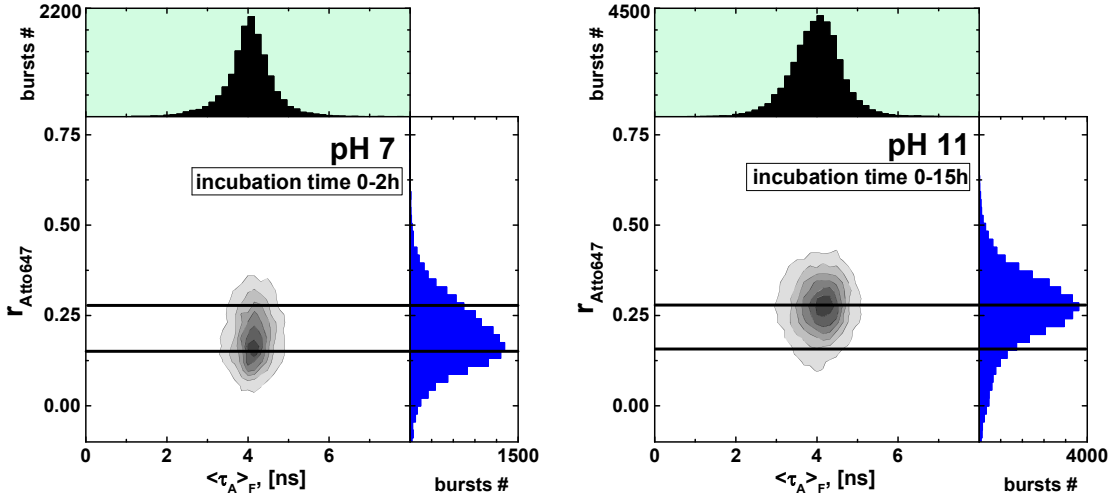
Supplementary Table 4: Biotin kinetics. Additional measurements testing the influence of Biotin on the kinetic behavior of the toxins. Different concentrations of biotin and different fluorophores were tested. The quality/measurability of the samples is expressed as active molecules, meaning the ratio of molecules that changed their state/polarization. Using biotinylated sample resulted in an instant (not resolvable) and fast (within one data point) change of PrePore to Pore state. Slower kinetics were fitted using a single exponential term and resulting tau is given.

Sample	Dye	Biotin	Active molecules	tau [h]
TcA-1193	Atto647N	none	2%	6.5±6
TcA-1193	BFL	none	23%	3h
TcA-1193	Atto647N	none	6%	5±1.5
TcA-1193	Atto647N	0.5x	5%	7h±5
TcA-1193	Atto647N	2x	15%	fast
Tca-1193-QTY	Atto647N	2x	20%	instant
TcA-1279	Atto647N	2x	37%	fast
TcA-1193	Atto647N	50x	0%	
TcA-2365	BFL	none	29%	20h±3
TcA -2384	BFL	none	17%	20h±3
TcA-2365	Atto647N	none	11%	6h±3
TcA-2365	Atto647N	2x	11%	fast, 3h±1.5

a – TcA(914-BFL)-Biotin:TcB(1041-Atto647N)-TcC, hetero FRET



b – TcA(1279-Atto647N)-Biotin , homo FRET



Supplementary Figure 9: Multidimensional smFRET histograms of TcA(914-BFL)-Biotin:TcB-TcC(1041-Atto647) and TcA(1279-Atto647N)-Biotin. a Two-dimensional histogram FRET efficiency *E* vs lifetime of donor in presence of acceptor $\langle\tau_{D(A)}\rangle_f$. One-dimensional histograms are the projected burst distributions over a single variable. Acceptor and donor only bursts were filtered out using stoichiometry and Alex 2CDE filter. FRET efficiency values are corrected for background, spectral crosstalk, direct acceptor excitation and detection efficiency ratio. Static FRET line (blue line) was calculated using [33] $E(\langle\tau_{D(A)}\rangle_f)=1-(((0.0030*\langle\tau_{D(A)}\rangle_f^4)+(-0.0535*\langle\tau_{D(A)}\rangle_f^3)+(0.3118*\langle\tau_{D(A)}\rangle_f^2)+0.4093*\langle\tau_{D(A)}\rangle_f^1-0.0270)/\langle\tau_{D(0)}\rangle_f)$. b Two-dimensional histogram of scatter corrected anisotropy of and lifetime $\langle\tau_A\rangle_f$ of TcA(1279-Atto647N)-Biotin. Anisotropy values are corrected for background. Horizontal lines are showing the mean value of anisotropy under pH-7 environment (r = 0.15) and pH-11 environment (r = 0.27).

Supplementary Table 5: Distance and efficiency predictions for the pore injection sample TcA(914-BFL:TcB-TcC(1041-Atto647)

		subunit				
state		**A**	**B**	**C**	**D**	**E**
PrePore	E_{FRET}	0.00	0.00	0.00	0.00	0.00
	$\langle R_{DA} \rangle_E$	163 Å	179 Å	202 Å	201 Å	180 Å
Pore	E_{FRET}	0.25	0.01	0.00	0.00	0.02
	$\langle R_{DA} \rangle_E$	58 Å	105 Å	147 Å	143 Å	95 Å

R_0=49 Å

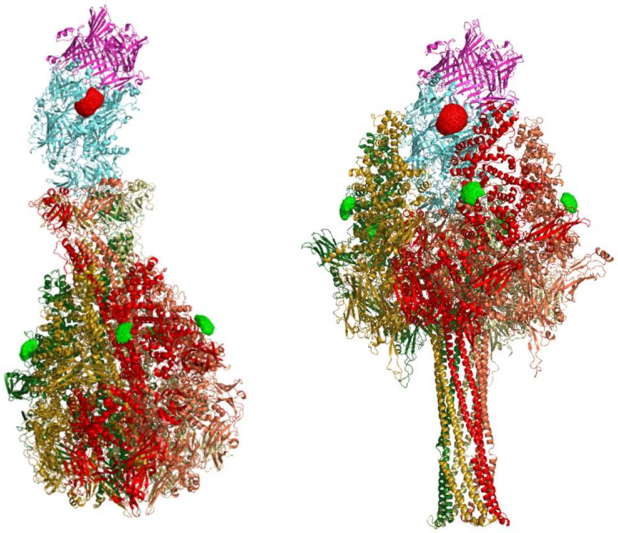
Supplementary Table 6: Distance and efficiency predictions for the shell destabilization sample TcA(1279-Atto647N)

state		**A- B**	**A- C**	**A- D**	**A-E**
PrePore	E_{FRET}	0.85	0.26	0.26	0.84
	$\langle R_{DA} \rangle_E$	50 Å	80 Å	80 Å	51 Å
Pore	E_{FRET}	0.19	0.01	0.01	0.19
	$\langle R_{DA} \rangle_E$	86 Å	139 Å	138Å	85 Å

R_0=67 Å

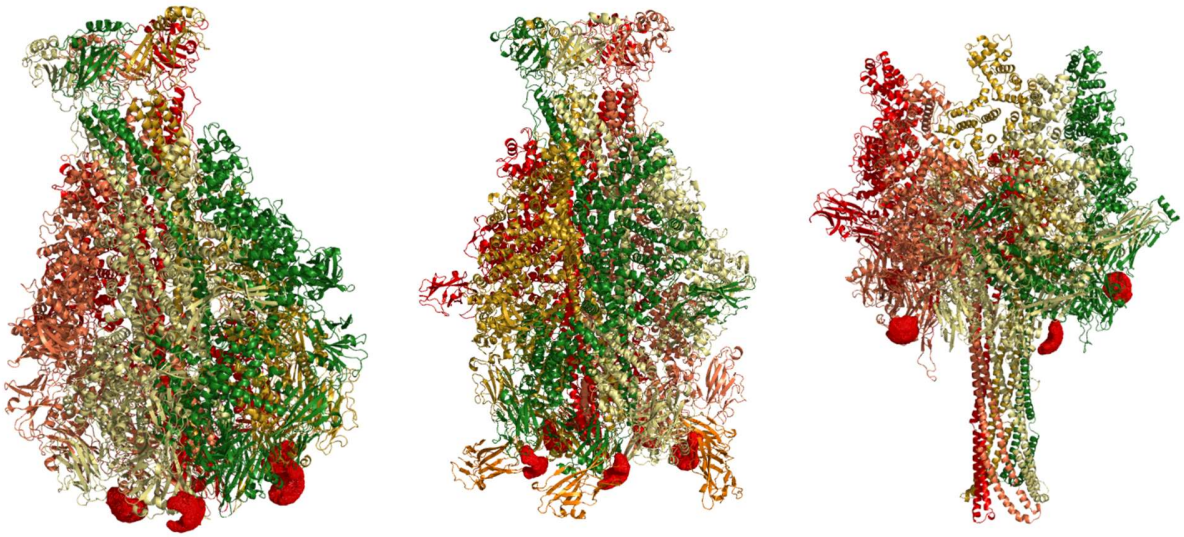
a – Pore Injection TcA(914-BFL:TcB(1041-Atto647N)-TcC-Biotin

PrePore **Pore**

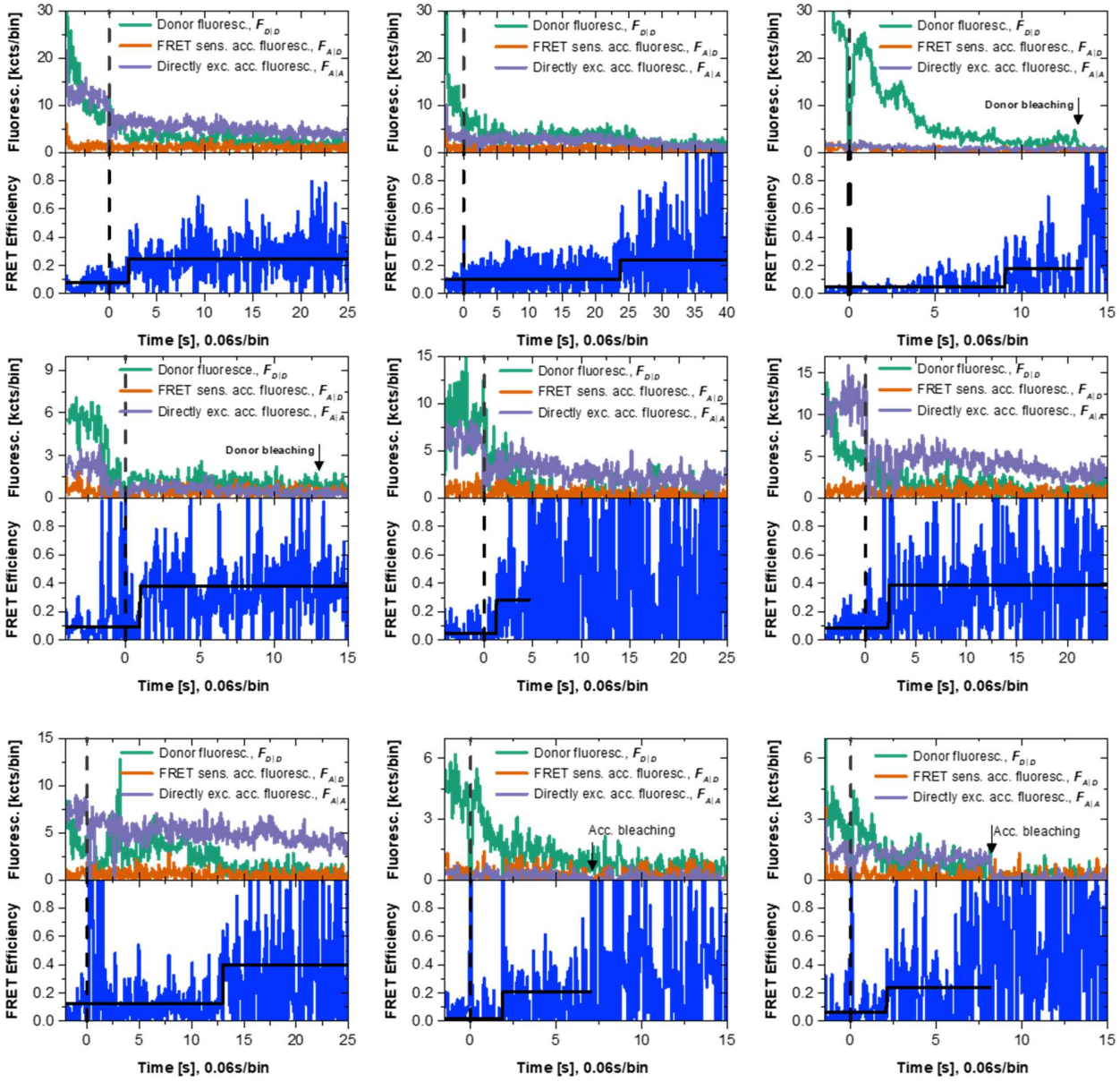


b – Shell destabilization TcA(1279-Atto647N)-Biotin

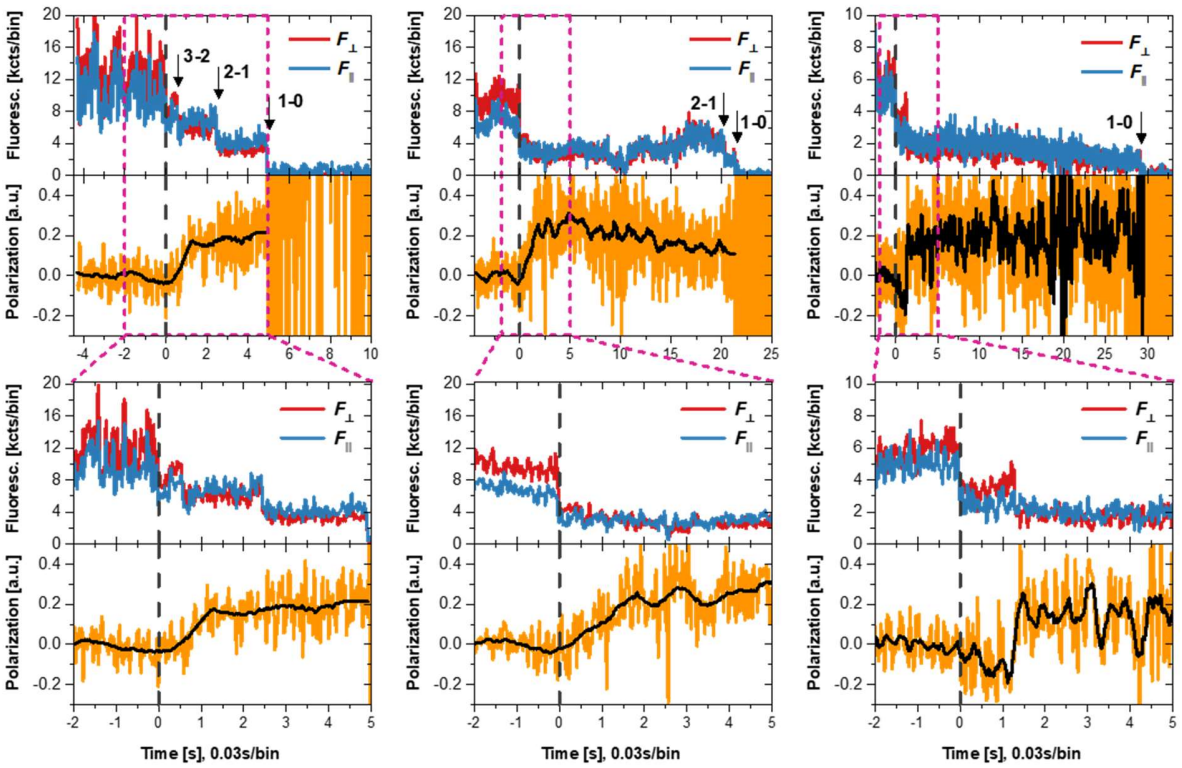
PrePore **Intermediate** **Pore**

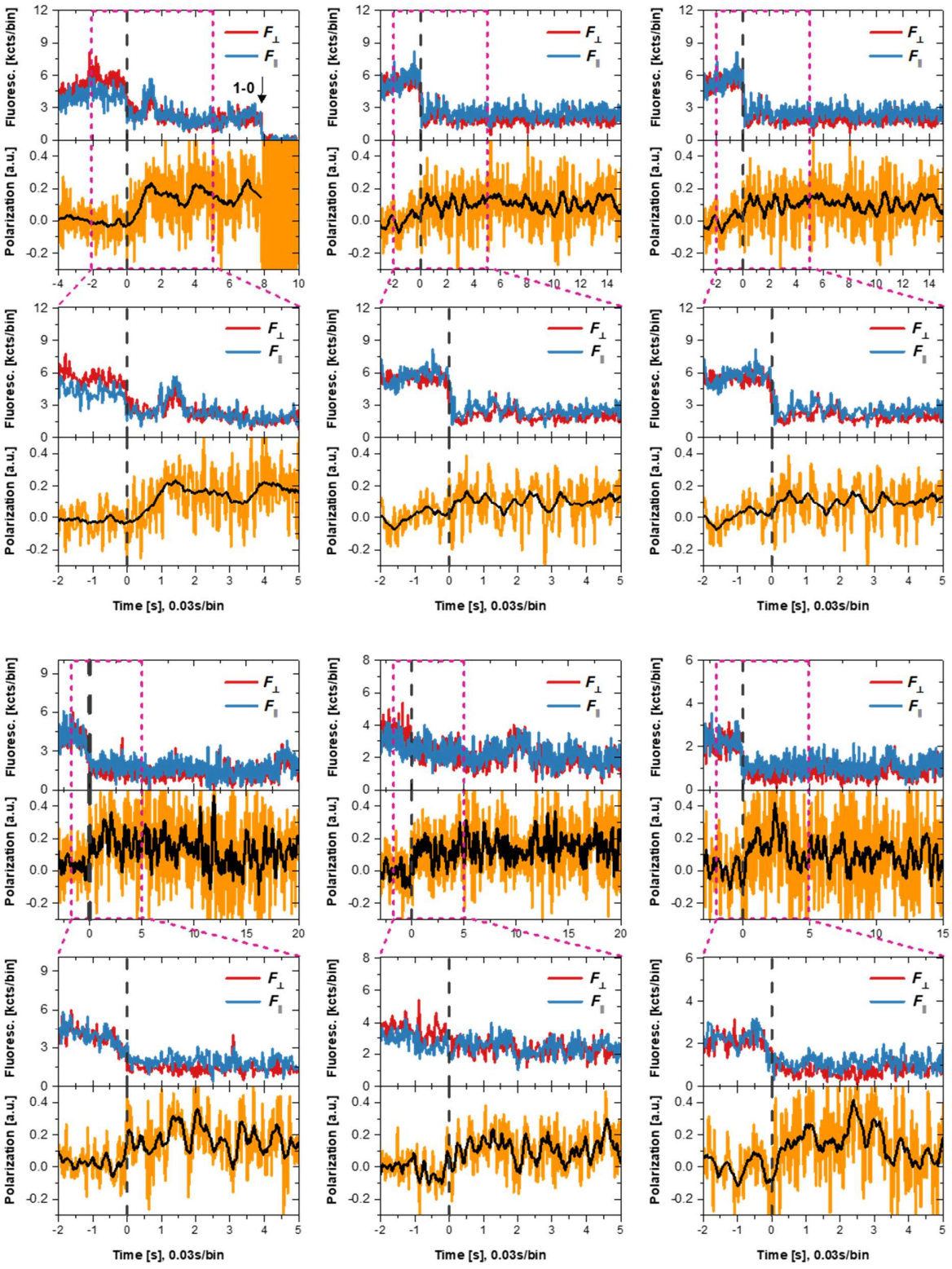


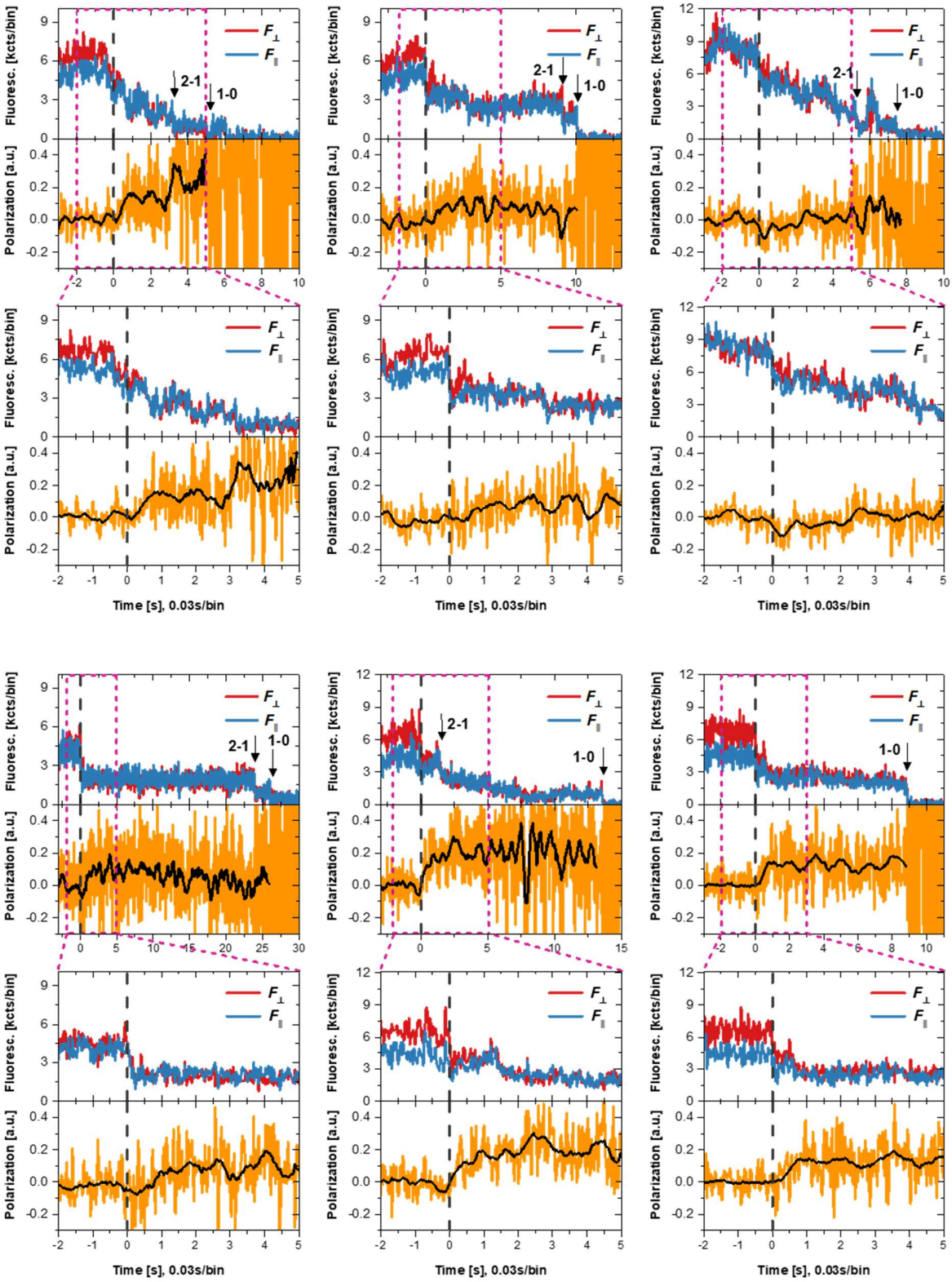
Supplementary Figure 10: Structures for pore injection and shell destabilization. a 3D structures with Accessible Volume (AV) of BFL at TcA914 and Atto647N at TcB-TcC(1041). **b** 3D-structures from PrePore to intermediate to Pore with AV of Atto647N attached to TcA1279

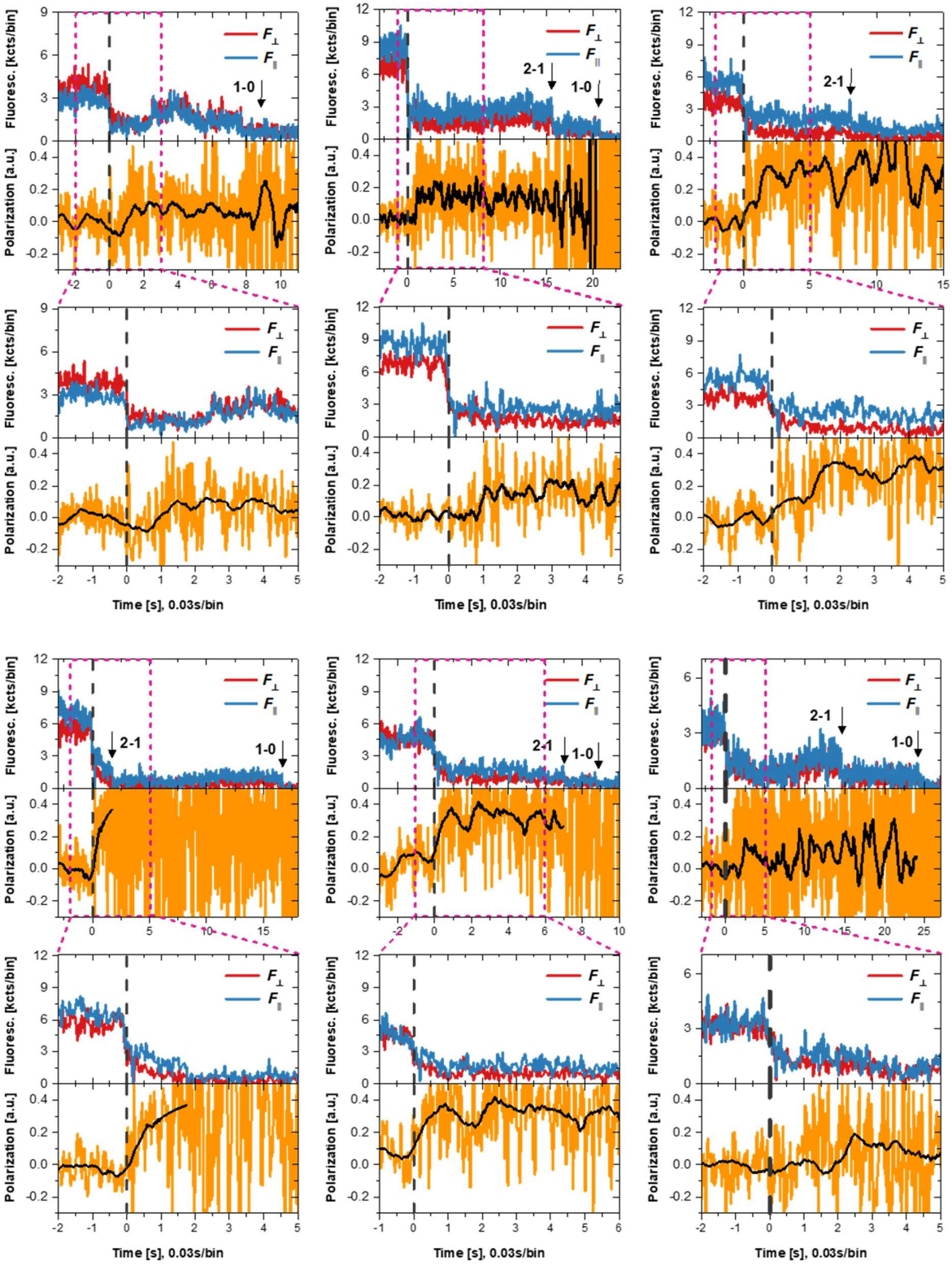


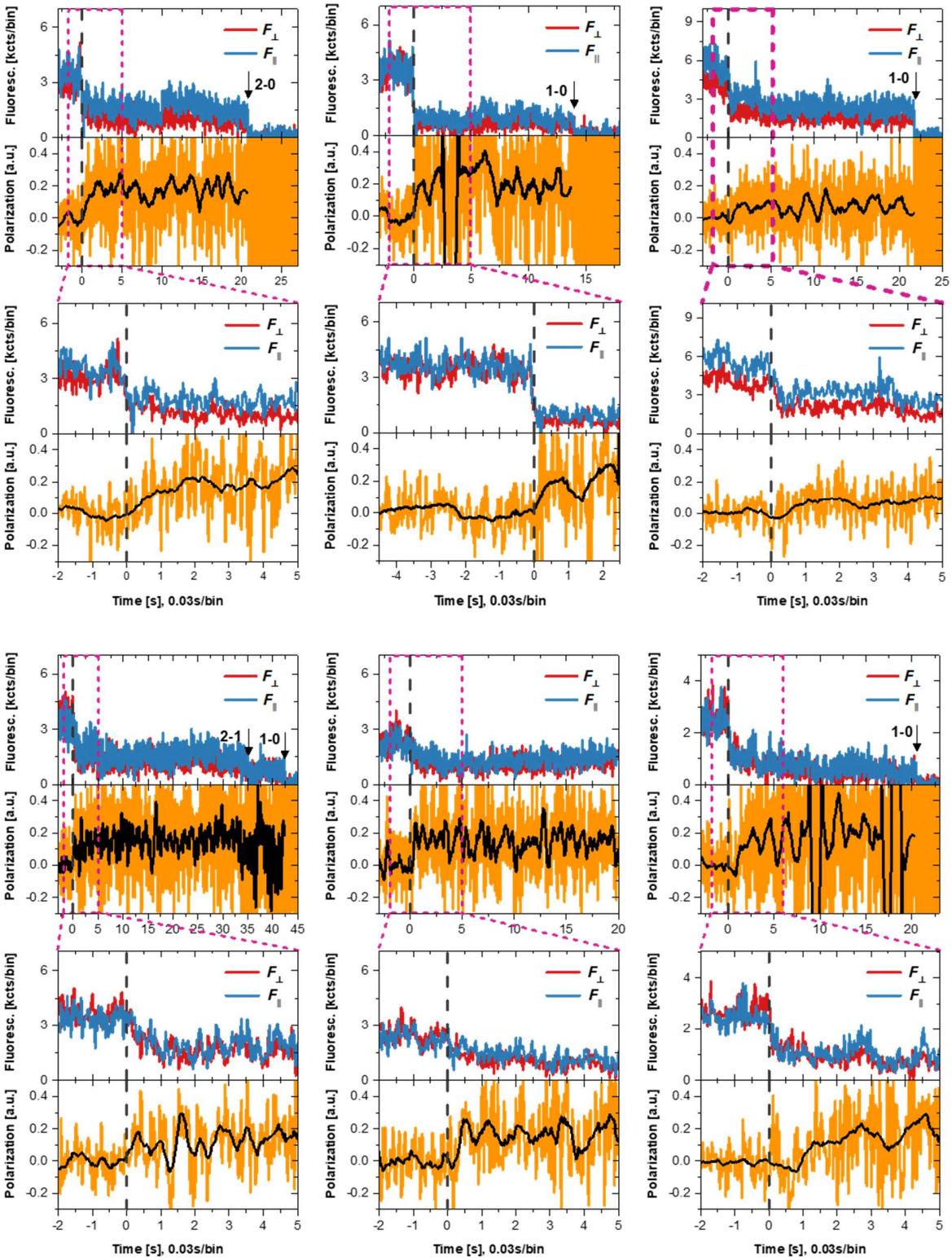
Supplementary Figure 11. Measured traces of pore injection. The sample TcA(914-BFL)-Biotin:TcB(1041-Atto647N)-TcC was labeled with a ratio of one BFL per TcA-pentamer so that heteroFRET efficiency for a single dye pair was measured. Top layer is displaying the background, crosstalk and detection ratio corrected fluorescence of the donor, $F_{D|D}$, the FRET sensitized acceptor, $F_{A|D}$, and of the directly excited acceptor $F_{A|A}$. The black line in the bottom layer is showing the average efficiency value before and after the transition, approximately matching the expected efficiencies calculated using 3D structures of PrePore/Pore state (see Supplementary Table 5). The dashed black line is indicating the pH change, offset corrected to t=0. After a relaxation time the jump is within a single bin (bin width=58.88 ms). Before and after the transition the FRET efficiency stays at a constant average value until bleaching occurs

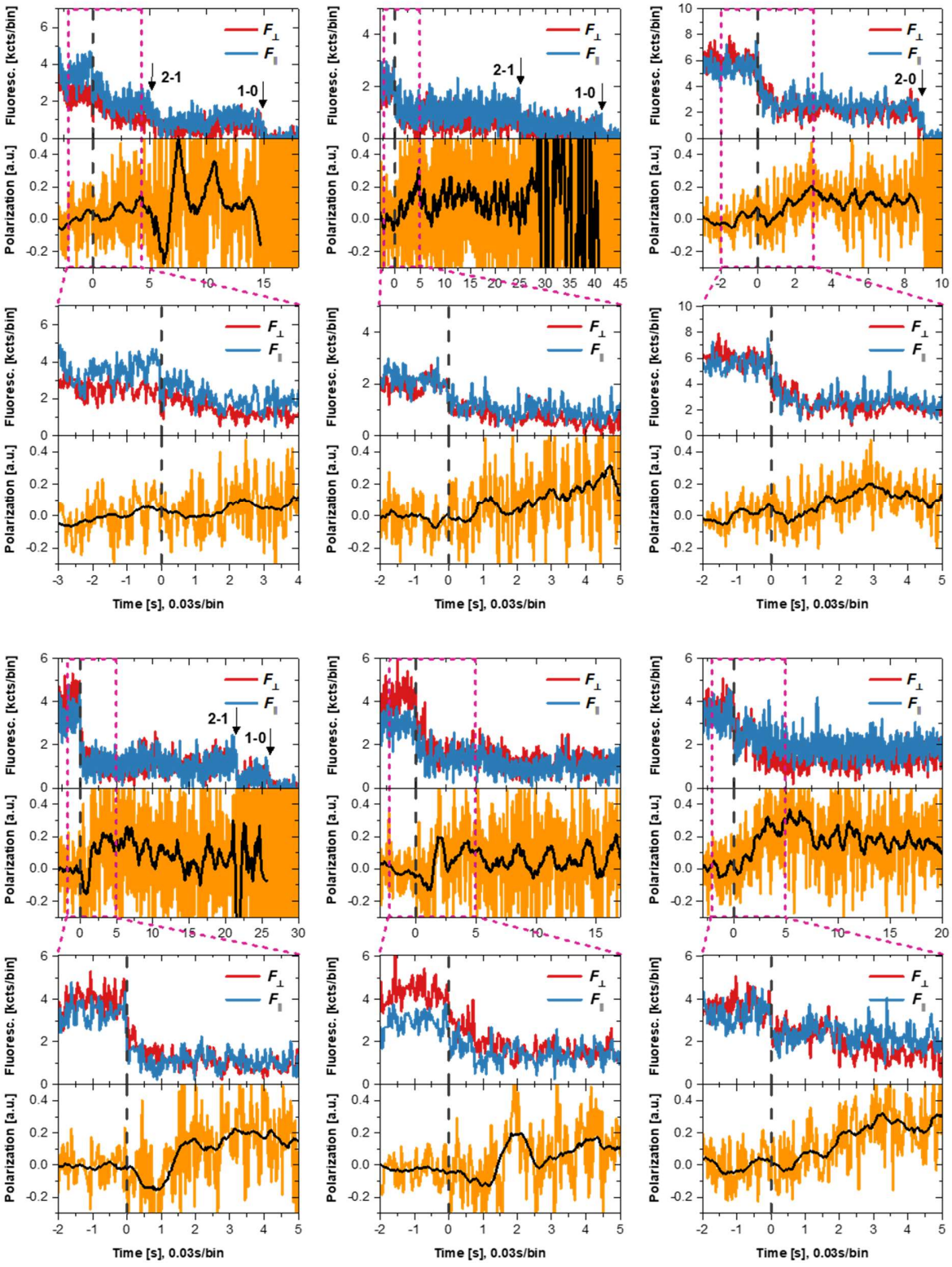

Fluoresc. [kcts/bin]
Polarization [a.u.]
$F_\perp$
$F_\parallel$
3-2
2-1
1-0
Time [s], 0.03s/bin

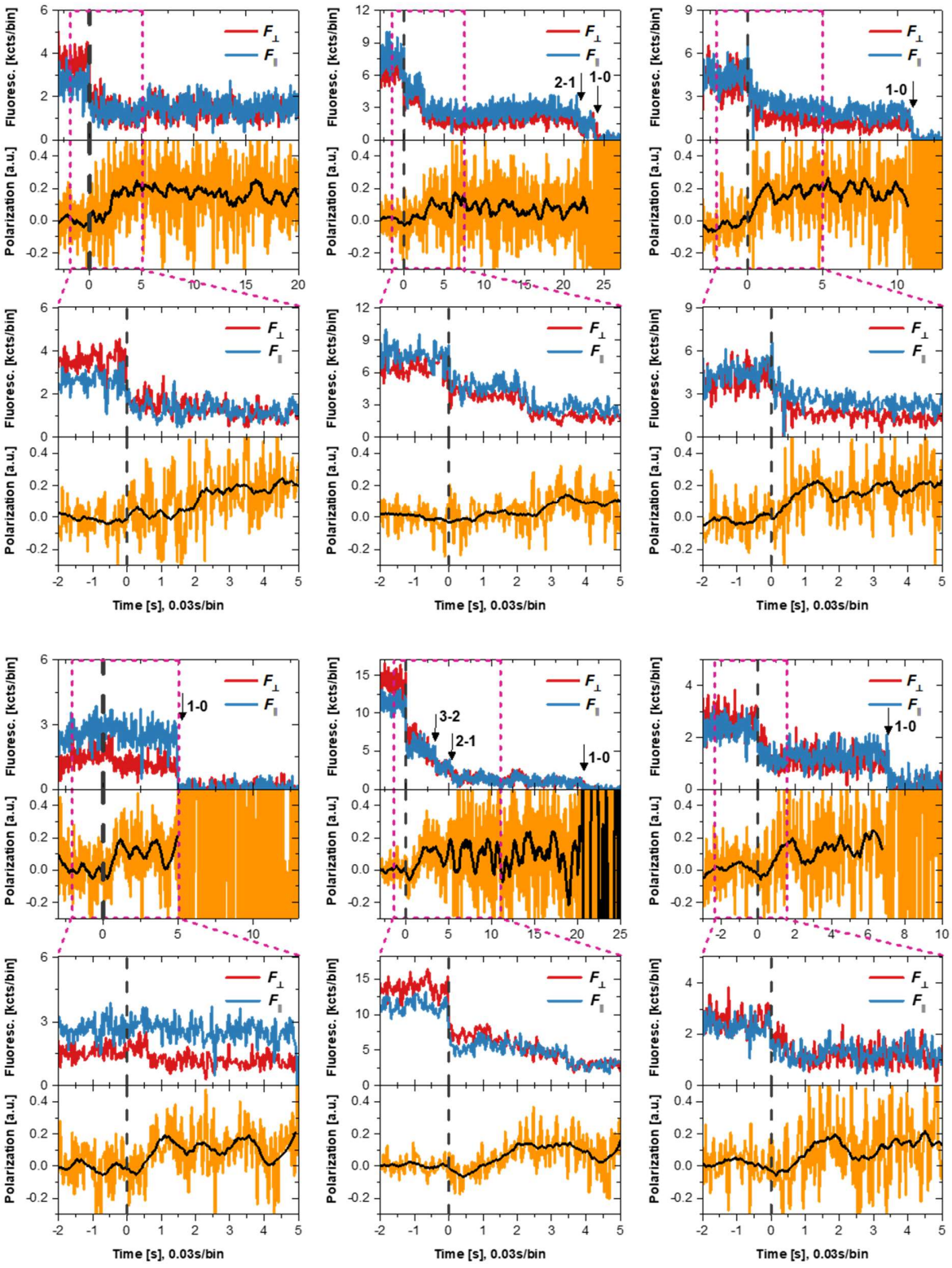

Fluoresc. [kcts/bin]
Polarization [a.u.]
$F_\perp$
$F_\parallel$
1-0
Time [s], 0.03s/bin

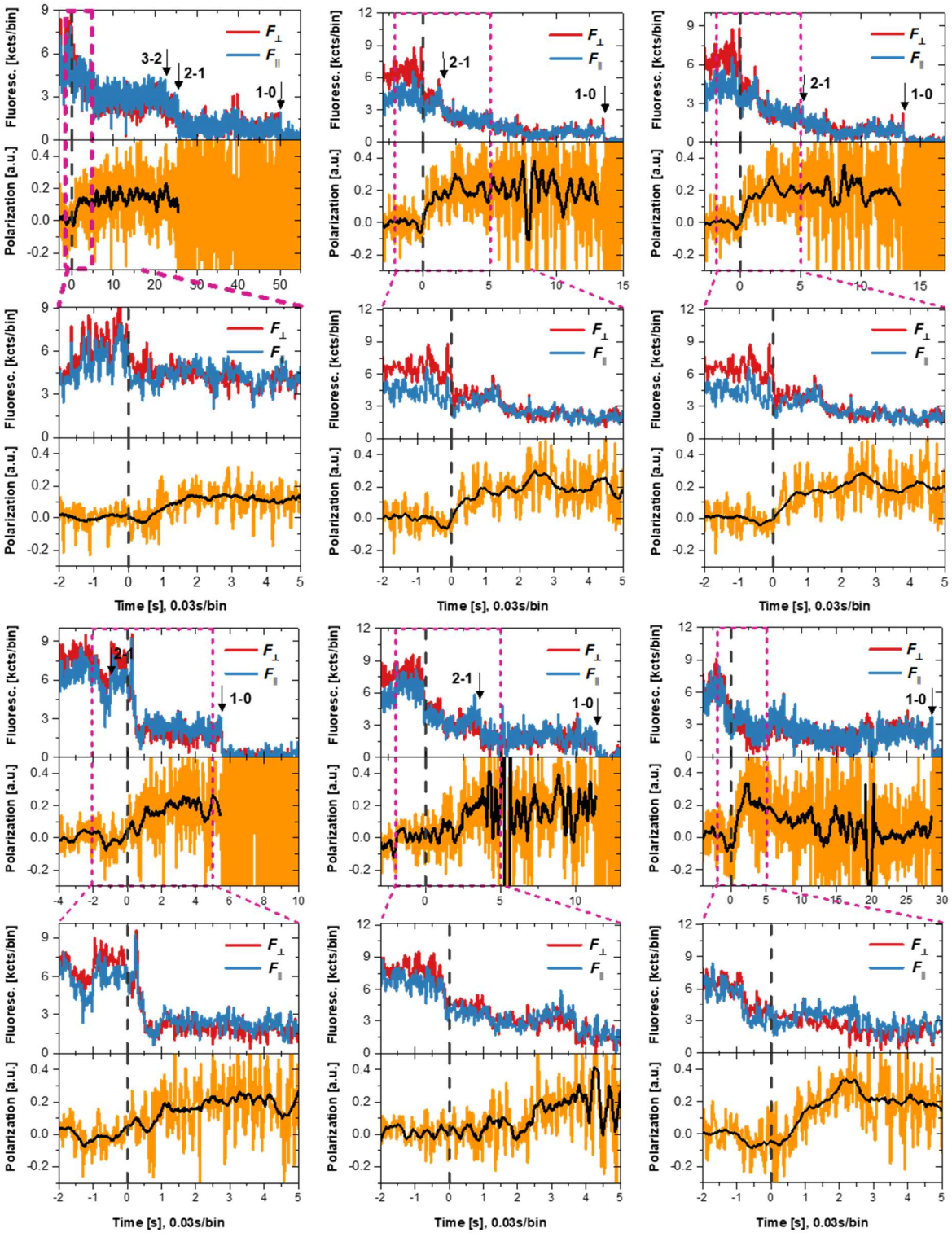

Fluoresc. [kcts/bin]
Polarization [a.u.]
$F_\perp$
$F_\parallel$
2-1
1-0
Time [s], 0.03s/bin

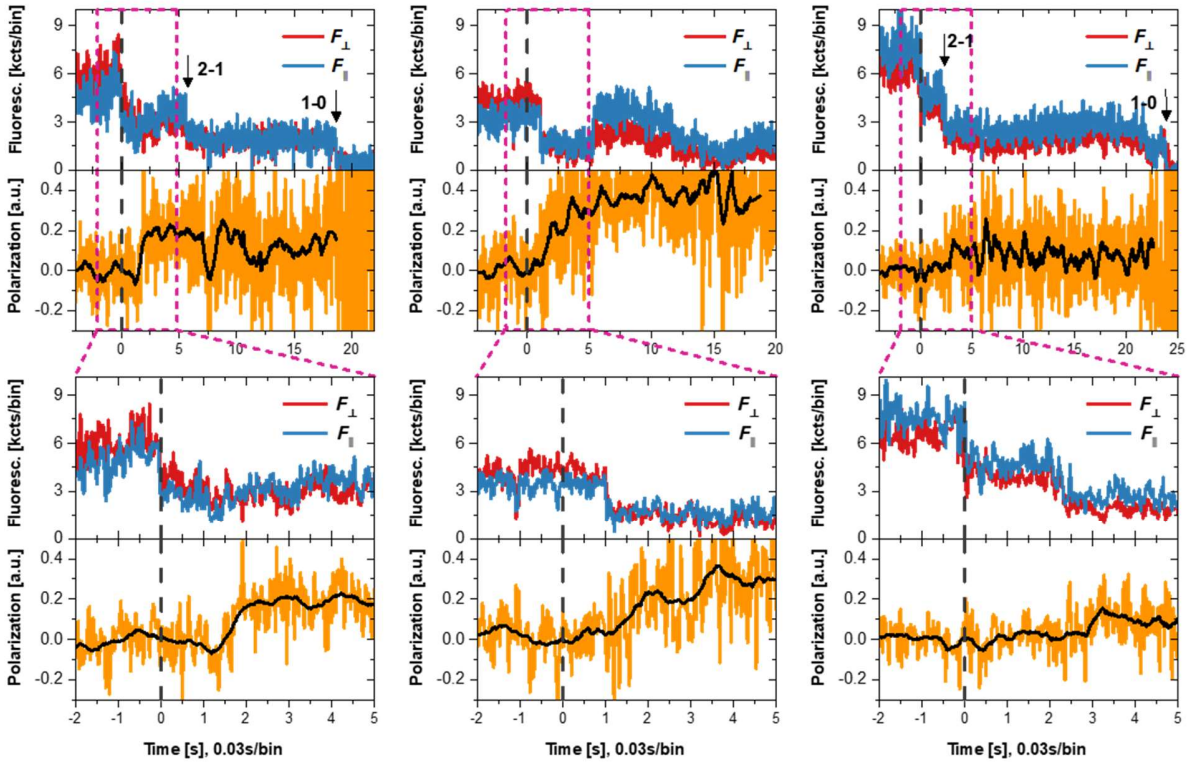

Fluoresc. [kcts/bin]
Polarization [a.u.]
Time [s], 0.03s/bin
F⊥
F∥
1-0
2-1


Fluoresc. [kcts/bin]
Polarization [a.u.]
$F_\perp$
$F_\parallel$
2-0
1-0
2-1
Time [s], 0.03s/bin


Fluoresc. [kcts/bin]
Polarization [a.u.]
$F_\perp$
$F_\parallel$
2-1
1-0
2-0
Time [s], 0.03s/bin


Fluoresc. [kcts/bin]
Polarization [a.u.]
Time [s], 0.03s/bin
F⊥
F∥
1-0
2-1
3-2


Fluoresc. [kcts/bin]
Polarization [a.u.]
$F_\perp$
$F_\parallel$
3-2
2-1
1-0
Time [s], 0.03s/bin



Supplementary Figure 12: Measured traces of the shell destabilization using an anisotropy assay that senses local environment and homoFRET. TcA(1279-Atto647N)-Biotin was labelled with a ratio of 4 dyes per pentamer. Top layer is displaying the background corrected perpendicular polarized Fluorescence $F_\perp$, and the parallel polarized Fluorescence $F_\parallel$. Dashed magenta lines are indicating the zoom interval for the lower diagram. Arrows mark the intensity drop due to bleaching with the numbers of fluorophores attached to the toxin in bright state before and after the bleaching event. Bottom layer is showing the polarization value (orange) and the smoothed polarization trajectory using the Savitzky-Golay algorithm (black).

7 Manuscripts

7.1 Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distance in proteins

Ganesh Agam[1,#], Christian Gebhardt[2,#], Milana Popara[3,#], Rebecca Mächtel[2,], Julian Folz[3], [Benjamin Ambrose[4], Neharika Chamachi[5,6], Sang Yoon Chung[25], Timothy D. Craggs[4], Marijn de Boer[7], Dina Grohmann[8], Taekjip Ha[9], Andreas Hartmann[5,6], Jelle Hendrix[10,11], Verena Hirschfeld[12], Christian G. Hübner[12], Thorsten Hugel[13,14], Dominik Kammerer[15], Hyun-Seo Kang[16], Achillefs N. Kapanidis[15], Georg Krainer[5,6], Kevin Kramm[8], Edward Lemke[18,19], Eitan Lerner[20], Emmanuel Margeat[21], Kirsten Martens[22], Jens Michaelis[23], Jaba Mitra[9,24], Gustavo G. Moya Muñoz[2], Robert B. Quast[21], Nicole B. Robb[15,§], Michael Sattler[16,17], Michael Schlierf[5,6], Jonathan Schneider[2], Tim Schröder[1], Anna Sefer[23], Piau Siong Tan[18], Johann Thurn[13,‡], Philip Tinnefeld[1], John van Noort[22], Shimon Weiss[25,26], Nicolas Wendler[2] Niels Zijlstra[2]], Anders Barth[2,3,†,*], Claus A. M. Seidel[3,*], Don C. Lamb[1,*], Thorben Cordes[2,*]

[1] Department of Chemistry, Ludwig-Maximilians-Universitaet Muenchen, Butenandtstr. 5-13, 81377 München, Germany

[2] Physical and Synthetic Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, Groβhadernerstr. 2-4, 82152 Planegg-Martinsried, Germany

[3] Molecular Physical Chemistry, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany

[4] Department of Chemistry, University of Sheffield, S3 7HF, UK

[5] B CUBE – Center for Molecular Bioengineering, TU Dresden, Tatzberg 41, 01307 Dresden, Germany

[6] Cluster of Excellence Physics of Life, Technische Universität Dresden, 01062 Dresden, Germany

[7] Molecular Microscopy Research Group, Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

[8] Department of Biochemistry, Genetics and Microbiology, Institute of Microbiology, Single-Molecule Biochemistry Lab, University of Regensburg, Regensburg, Germany

[9] Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, and Howard Hughes Medical Institute, Baltimore, MD 21205

[10] Dynamic Bioimaging Lab, Advanced Optical Microscopy Center and Biomedical Research Institute, Hasselt University, Agoralaan C (BIOMED), B-3590 Hasselt, Belgium

[11] Department of Chemistry, KU Leuven, Celestijnenlaan 200F, B-3001 Leuven, Belgium

[12] Institute of Physics, University of Lübeck, Germany

[13] Institute of Physical Chemistry, University of Freiburg, Germany

[14] Signalling Research Centers BIOSS and CIBSS, University of Freiburg, Schänzlestrasse 18, 79104 Freiburg, Germany

[15] Department of Physics, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK.

Kavli Institute of Nanoscience Discovery, University of Oxford, Oxford OX1 3QU, UK

[16] Bayerisches NMR Zentrum, Department Chemie, Technische Universität München, School of Natural Sciences, Lichtenbergstr. 4, DE-85747 Garching, Germany

[17] Institute of Structural Biology, Helmholtz Center Munich, Ingolstaedter Landstr. 1, 85764 Neuherberg Biomolecular NMR Spectroscopy

[18] Biocenter, Johannes Gutenberg University Mainz; Mainz, 55128, Germany.

Institute of Molecular Biology; Mainz, 55128, Germany

[19] Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany

[20] Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, and The Center for Nanoscience and Nanotechnology, Faculty of Mathematics & Science, The Edmond J. Safra Campus, The Hebrew University of Jerusalem, Jerusalem, 91904, Israel

[21] Centre de Biologie Structurale (CBS), Univ. Montpellier, CNRS, INSERM, Montpellier, France

[22] Biological and Soft Matter Physics, Huygens–Kamerlingh Onnes Laboratory, Leiden University, Leiden, The Netherlands

[23] Institute for Biophysics, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany

[24] Materials Science and Engineering, University of Illinois Urbana-Champaign, Urbana IL 61801, USA;

[25] Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA

[26] California NanoSystems Institute, University of California, Los Angeles, CA 90095, USA

§ current address: Warwick Medical School, The University of Warwick, Coventry, CV4 7AL, UK

‡ current address: German Aerospace Center (DLR), Institute for Technical Physics, Solid-state Laser and Nonlinear Optics, Pfaffenwaldring 38-40, 70569 Stuttgart

† current address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands.

# these authors contributed equally to this paper

* corresponding author emails: anders.barth@hhu.de, d.lamb@lmu.de, cseidel@hhu.de, cordes@bio.lmu.de

[] authors within brackets are listed alphabetically

Abstract

Single-molecule FRET (smFRET) has become an established tool to study biomolecular structure and dynamics *in vitro* and in live cells. We performed an international blind study involving 19 labs to assess the uncertainty of FRET experiments for proteins with respect to the measured FRET efficiency histograms, determination of distances, and the detection and quantification of structural dynamics. Using two protein systems that undergo distinct conformational changes, we obtained an uncertainty of the FRET efficiency of less than ± 0.06, corresponding to an interdye distance precision of ≤ 0.2 nm and accuracy of ≤ 0.5 nm. We further discuss the limits for detecting distance fluctuations with sensitivity down to ≲ 10% of the Förster distance and provide guidelines on how to detect potential dye perturbations. The ability of smFRET experiments to simultaneously measure distances and avoid averaging of conformational dynamics slower than the fluorescence lifetime is unique for dynamic structural biology.

Introduction

Förster resonance energy transfer (FRET) studies have become a mature and widely-used approach that is complementary to classical structural biology techniques1–4. FRET experiments provide information on the structure and conformational heterogeneity of biomolecules over a distance range of 3 to 12 nm in space and, when performed on single-molecules, can provide additional information regarding conformational dynamics on the timescales of nanoseconds to seconds[1,2,5–10]. It allows for the quantitative assessment of structural dynamics and the heterogeneity of conformational ensembles, which are not easily accessible by x-ray crystallography, cryo-electron microscopy and techniques such as cross-linking mass-spectrometry, that provide structural information of solution structures, but lack temporal information. It can also be used to resolve parts of structures or even full structures of biomolecules in an integrative manner (for examples, see refs [11–17]) and has the unique ability to provide correlated information on structure and dynamics[1,2].

Hellenkamp *et al.* presented a quantitative multi-laboratory single-molecule FRET (smFRET) blind study of 20 different laboratories using static double-stranded DNA (dsDNA) oligonucleotide rulers that demonstrated a high reproducibility between the different labs with an uncertainty of less than 6 Å for the FRET-derived distances[18]. Although the optimal procedure for determining correction factors involved in converting setup-dependent FRET efficiency values into accurate distances remains a topic of active discussion[1], the results presented by Hellenkamp *et al.* strongly support the idea that standardized smFRET measurements are a useful addition for integrative modelling of static biomolecular structures[12,19,20].

Here, we take the next step by assessing whether the established procedures translate to more flexible biomacromolecules such as proteins, which often undergo conformational fluctuations. Compared to dsDNA, proteins are generally more challenging systems to study because the local chemical environments of the tethered dyes can vary significantly, which is further amplified by conformational dynamics. Site-specific labeling of proteins usually requires the introduction of point mutations (typically cysteines), which can pose further complications with respect to the labeling procedures depending on the structure and nature of the protein1. Moreover, protein samples require careful handling and storage, due to sample instability and aggregation, and their sensitivity to the biochemical environment and experimental conditions such as buffer composition, pH, temperature or interaction with surfaces. In a blind study involving 19 labs, we investigated how reliably smFRET efficiency histograms of diffusing proteins can be measured by confocal detection of freely-diffusing molecules, and how well structural dynamics can be detected and quantified. As realistic and challenging test cases, we chose two proteins, the maltose-binding protein (MalE) and the U2 Auxiliary Factor 2 (U2AF2), which display conformational dynamics on different timescales that are modulated by ligand binding. Fluorescently labeled protein samples were prepared by stochastically labeling protein double-cysteine variants at positions that will report on specific intramolecular distances. Two key questions are addressed here: (i) How consistently can smFRET efficiency histograms (and the derived distances) be determined by different labs for protein samples? (ii) How reliably can smFRET measurements detect and quantify structural dynamics in proteins? In this context, we investigated the minimal structural fluctuations detectible by smFRET measurements and discuss how to achieve this sensitivity.

Our comparison study confirmed the reproducibility of measuring accurate FRET efficiency histograms and the ability of smFRET to detect and quantify conformational dynamics on the sub-millisecond timescale. We demonstrate reproducible FRET efficiency values with uncertainties of less than ± 0.06 corresponding to a distance precision of ≤ 2 Å and an accuracy ≤ 5 Å in MalE. Moreover, we compare the variability of setup-dependent detection parameters and characterize the calibration uncertainty. To push the detection limits for structural dynamics, we refined established experimental and data analysis procedures for the characterization of dynamics and studied a series of distinct dye pairs to identify and eliminate dye-specific effects. With this refinement, we could detect distance fluctuations on the order of 5 Å in the FRET sensitive range. Our work demonstrates the capability of smFRET experiments to study challenging and realistic protein systems with conformational dynamics on timescales from nanoseconds to seconds, highlighting their importance in the expanding toolbox of dynamic integrative structural biology[19–21].

Results

In this study, we chose two prototypic protein systems that exhibit conformational dynamics on different timescales. Our first target was MalE of *E. coli.* It is a periplasmic component of the ATP binding cassette transporter $MalFGK_2$-E[22,23] and has been widely studied and applied in biochemistry and molecular biology[24]. MalE exhibits a type II periplasmic-binding protein fold[25,26] composed of two rigid domains connected by a flexible two-segment hinge (Fig. 1a). This domain arrangement enables an allosterically-driven motion from an open to closed state upon maltose binding with conformational dynamics on the sub-second timescale (Supplementary Fig. 1). As a second system, we chose the large subunit of the U2 auxiliary factor (U2AF2) of the pre-mRNA splicing machinery (spliceosome)[27]. The two RNA recognition motif domains (RRM1,2) of U2AF2 are connected by a long flexible linker and bind single-stranded Py-tract RNA with an affinity of K_d ~1.3 µM for the U9 RNA used in this study[28]. For U2AF2, the two domains fluctuate between an ensemble of detached conformations and a compact conformation in the apo state[29], whereas ligand binding stabilizes an open conformation (Fig. 2a)[30]. To avoid additional complexity and to restrict any pre-knowledge regarding the samples, the proteins were labeled and checked for functionality before being delivered to the participants.

SmFRET experiments were blindly performed by 19 laboratories for MalE and by seven laboratories for U2AF2 using different implementations of solution-based confocal spectroscopy with alternating excitation, µs-ALEX[31] for intensity-based analysis and ns-ALEX[32] or PIE[33] for intensity- and lifetime-based analyses (Supplementary Fig. 2). We adapted a data analysis routine similar to that of Hellenkamp *et al.*[18] to determine setup-independent accurate FRET efficiency *E* values from the photon counts detected in the donor (D) and acceptor (A) detection channels during a single-molecule event. The implementation of ALEX or PIE[2,8,9,18,34,35] (see Supplementary Note 1 for a detailed comparison of ALEX and PIE) was crucial for: (i) careful corrections of the registered photon counts to reflect the actual donor and acceptor signal; and (ii) exclusion of single-molecule events from further analysis that originate from incompletely labeled molecules, or showed photo-blinking or bleaching. The correction procedures for reporting accurate FRET efficiencies are described in the Online Methods and include subtraction of background signal

from all channels and the determination of four correction factors: (α) for spectral crosstalk of D fluorescence into the A channel, (β) for normalization of direct D and A excitation fluxes, (γ) for differences in donor and acceptor quantum yields and detection efficiencies, and (δ) for the ratio of indirect and direct A excitation (Supplementary Fig. 3 and Supplementary Tables 1 and 2)[34].

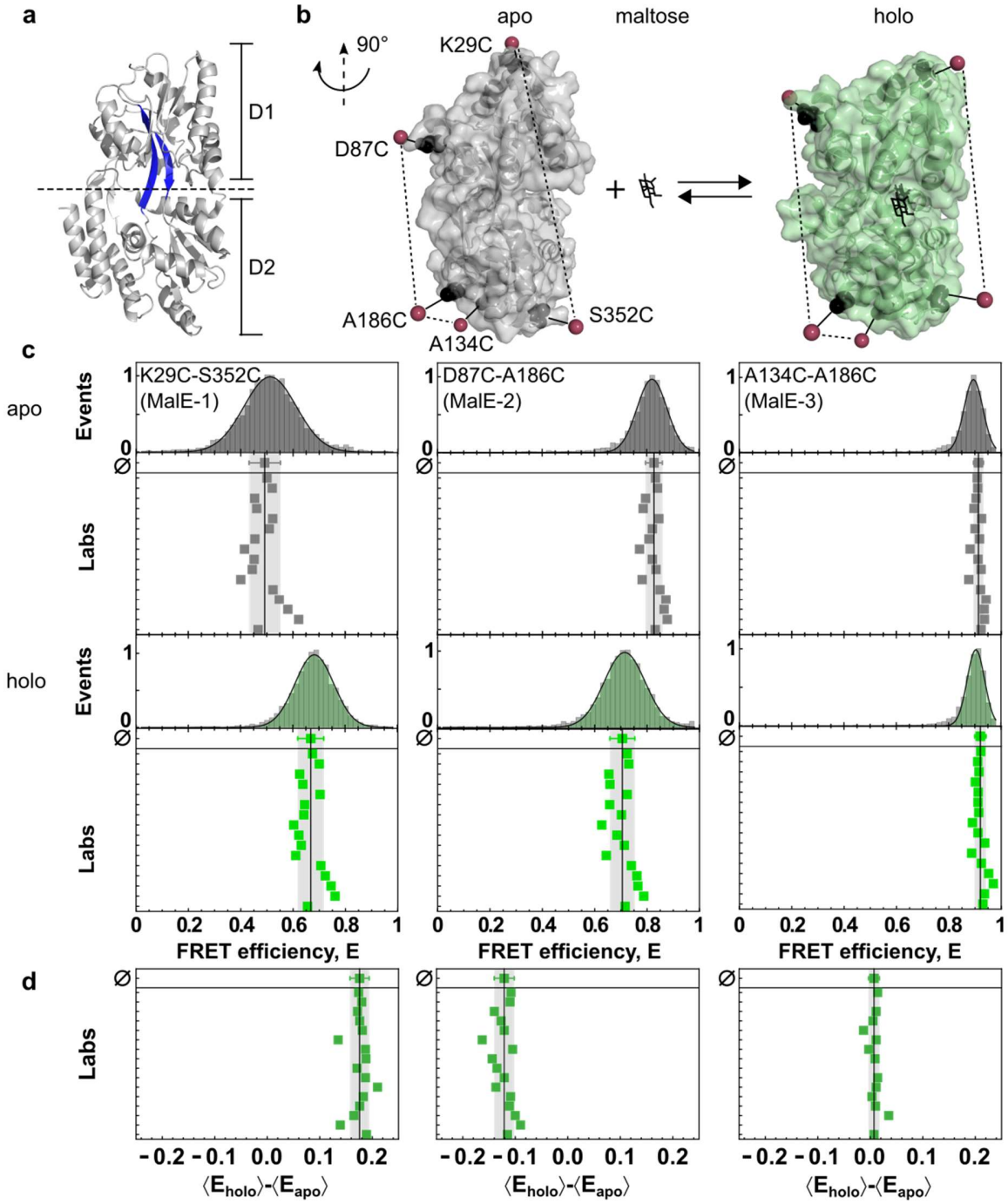


Fig. 1. Experimental design of MalE as a protein model system for smFRET studies: (a) Crystal structure of MalE in its ligand-free apo state (PDB-ID: 1OMP) with domains D1 and D2 linked by flexible beta-sheets (highlighted in blue). **(b)** The crystal structure of MalE (rotated by 90° as compared to **a**) in the apo (grey, PDB-ID: 1OMP) and holo (green, PDB-ID: 1ANF) states with mutations at K29C / S352C (MalE-1), D87C / A186C (MalE-2), and A134C / A186C (MalE-3) indicated in black. Note, each mutant only contains one cysteine pair and were measured using the Alexa546-Alexa647 FRET pair. The estimated mean position of the fluorophores from AV calculations are shown as red spheres. **(c)** FRET efficiency *E* histograms for three MalE mutants, MalE-1 (left), MalE-2 (middle), and MalE-3 (right), in the absence and presence of 1 mM maltose (bottom, green) for one exemplary dataset measured in lab 1. The distribution is fitted to a Gaussian distribution. The reported mean FRET efficiencies for 16 labs are shown below (due to experimental difficulties, the results of three labs were excluded; see Supplementary Table 1). The mean FRET efficiency and the standard deviation of all 16 labs are given by the black line and grey area. **(d)** Individual FRET efficiency differences for each lab, between the apo and holo states, $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$, for MalE-1 (left), MalE-2 (middle), and MalE-3 (right).

MalE. For smFRET investigations of MalE, we prepared three different double-cysteine variants that cover a large part of the dynamic range of FRET and monitor the conformational change in the protein upon maltose binding (Fig. 1b, see Online methods and Supplementary Fig. 4). The variants were designed such that MalE-1 (K29C-S352C) shows a decrease in the inter-dye distance upon maltose binding, MalE-2 (D87C-A186C) shows an increase in distance and MalE-3 (A134C-A186C) shows no distance change upon substrate binding. All variants of MalE were stochastically labeled in one of the laboratories at the given positions with Alexa Fluor 546 (Alexa546) as the donor and Alexa Fluor 647 (Alexa647) as the acceptor fluorophore. Before shipment, we confirmed the functionality of the labeled protein by ligand titrations using smFRET and microscale thermophoresis and ensured that the ligand maltose does not affect the photophysical properties of the dyes (Supplementary Fig. 5 and 6). For the sake of comparison, participants were asked to provide the mean FRET efficiencies using the fit to a Gaussian distribution for estimating the peak of the apo and holo FRET efficiency histograms (as shown in Fig. 1c). For this study, we asked the participating laboratories to determine a common (global) γ-value using all three MalE for both the apo and holo measurement conditions (Supplementary Note 2 and Supplementary Fig. 3). To execute this workflow, participants used many different in-house or publicly available software packages following the given guidelines.

FRET efficiency histograms for representative experiments of the MalE variants in the apo state and in the presence of 1 mM maltose are shown in Fig. 1c together with values reported by 16 labs, showing very good agreement and reproducibility. It was not possible to extract accurate FRET efficiency values from three labs due to experimental difficulties such as a missing or suboptimal laser (see Supplementary Table 1 and Supplementary Note 2). All labs observed the expected maltose-induced conformational change for MalE-1 and MalE-2, and no significant change for MalE-3. This indicates that the samples did not degrade during shipment on dry ice and storage in the labs at 4 °C. MalE-1 showed an average FRET efficiency of 0.49±0.06 in the apo- and 0.67±0.05 in the holo state due to the hinge motion of the protein upon ligand binding. MalE-2 showed the expected decrease in FRET efficiency from 0.83±0.03 to 0.71±0.05 in the apo and holo states, respectively (Fig. 1c). MalE-3, with both labels on one lobe, showed no significant change in FRET efficiency (E_{apo} = 0.91±0.02, E_{holo} = 0.92±0.02).

The standard deviation of the determined mean FRET efficiency over all labs was less than ±0.06, similar to the precision found for dsDNA (Table 1 and Supplementary Table 3)[18]. We observe the highest standard deviation for MalE-1 and the lowest values of ±0.02 for MalE-3, which also has the highest FRET efficiency. As will be discussed in detail below, the observed spread of the reported FRET efficiencies depends less on the measurement statistics, but on the uncertainty in the calibration factors. This effect is largest at intermediate FRET efficiencies, which explains the higher spread of values for the MalE-1 mutant. Interestingly, for most labs, we observed systematic deviations of the reported FRET efficiency values for the apo and holo states from the mean value. This suggests that changes of the FRET efficiency are measured even more accurately than absolute values. In Fig. 1d, we analyze the individual FRET efficiency differences, $\langle E_{\mathrm{holo}}\rangle - \langle E_{\mathrm{apo}}\rangle$, between the apo and holo states for the different labs. Here, the distribution of values indeed narrows approximately by a factor of two for all three mutants because systematic deviations cancel out (standard deviations $\sigma_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$ for MalE-1: ±0.02, MalE-2: ±0.02, MalE-3: ±0.01, Fig. 1d, Table 1 and Supplementary Table 3).

Table 1. Average of mean FRET efficiency and standard deviation for MalE and U2AF2 samples reported by the participating laboratories. The calculated average $\mu_{\langle E\rangle}$ and standard deviation $\sigma_{\langle E\rangle}$ of the mean FRET efficiency values provided by the participating labs are given for all three studied mutants of MalE labeled with Alexa546 and Alexa647 under both apo and holo conditions (see Supplementary Table 3). The calculated mean and standard deviation of the difference in the reported mean FRET efficiency between the apo and holo ($\langle E_{\mathrm{holo}}\rangle - \langle E_{\mathrm{apo}}\rangle$) for the three MalE mutants are given by $\mu_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$ and $\sigma_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$ respectively (see Supplementary Table 3). The calculated average $\mu_{R_{\langle E\rangle}}$ and standard deviation $\sigma_{R_{\langle E\rangle}}$ of the mean distances were derived according to Eq. 2. The modeled distances $R_{\langle E\rangle}^{\mathrm{AV}}$ and $R_{\langle E\rangle}^{\mathrm{ACV}}$ are derived using accessible volume (AV) and accessible contact volume (ACV) calculations respectively, as described in the Online Methods. We also give the average and standard deviation for the FRET values determined for U2AF2 labeled with Atto532-Atto643 under both apo and holo conditions (Supplementary Table 4). *Only studied by two labs. **Due to the fast-structural dynamics in the sample, only 7 labs studied this mutant and distances were not determined. *** Only the holo state under holo condition was considered.

		Experimental Values			Modeled Distances	
Sample	**Condition**	$\mu_{\langle E\rangle} \pm \sigma_{\langle E\rangle}$	$\mu_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle} \pm \sigma_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$	$\mu_{R_{\langle E\rangle}} \pm \sigma_{R_{\langle E\rangle}}$ [Å]	$R_{\langle E\rangle}^{\mathrm{AV}}$ [Å]	$R_{\langle E\rangle}^{\mathrm{ACV}}$ [Å]
MalE-1	apo	0.49±0.06	0.177±0.019	65.4±2.6	72.0	67.7
	holo	0.67±0.05		57.8±2.1	62.1	58.3
MalE-2	apo	0.83±0.03	-	50.0±1.9	50.1	48.8
	holo	0.71±0.05	0.121±0.019	56.1±2.1	56.5	55.0
MalE-3	apo	0.913±0.019	0.007±0.010	43.8±1.7	39.9	38.9
	holo	0.920±0.021		43.0±2.4	40.8	39.8
MalE-4*	apo	0.442±0.025		67.6±1.2	67.8	64.3
	holo	0.678±0.017		57.4±0.7	56.9	54.6
MalE-5*	apo	0.613±0.003		60.22±0.15	61.8	59.3
	holo	0.821±0.001		50.43±0.08	49.3	48.2
U2AF2**	apo	0.74±0.03		49.6±1.3		
	holo***	0.46±0.04		60.8±1.7		

U2AF2. For the second protein, U2AF2, we chose the previously published double cysteine variant (L187C / G326C) of the minimal RRM1,2 construct, where we verified that protein properties of the used variant are not affected by labeling (Fig. 2a)[36,37]. The construct contains one cysteine on each RRM domain, which were labeled stochastically with the dye pair Atto532-Atto643. A subset of seven groups measured the second protein. To investigate the consistency of the shape of the obtained FRET efficiency histograms, we plot in Fig. 2b/c the smFRET histograms from the individual laboratories (row 1) as well as the average FRET efficiency distribution illustrated by the mean and standard deviation (row 2). All groups found a single broad distribution (Fig. 2b, row 1) in the apo state with an average FRET efficiency of E = 0.74±0.03 (row 2). In the presence of 5 µM ligand, a second narrower peak at lower E appears (Fig. 2c, row 1) with an average FRET efficiency of E= 0.46±0.04 (row 2) as expected for the open configuration of the holo state[30,36]. Notably, a significant fraction (~ 15%) of ligand-free protein remains in the sample at the RNA concentration used (Supplementary Fig. 7).

For the apo state, we obtained a similar standard deviation of ±0.03 as found for Mal-E, however a clear outlier was apparent (Supplementary Table 4). To test whether user bias affected the reported results, we had the datasets reanalyzed by a single person. While analyzing the different data sets, this person could determine an optimal procedure for determining the correction factors for this challenging sample (Supplementary Note 3). Hereby, the person could improve the agreement to a standard deviation of ±0.008 with no change in the average FRET efficiency value (Fig. 2d/e, Supplementary Table 4). The reanalysis revealed that the detection correction factor γ was the main cause of the deviations between the measurements as the single population of the apo state did not allow for a robust determination of the γ-factor[34,35]. In this case, it was best to estimate the γ-factor from a global analysis of the apo and holo measurements, which was possible due to the absence of any detected changes in the quantum yield of the fluorophores upon binding of the RNA (as measured using PIE) (Supplementary Table 5). We also reanalyzed data from the same seven laboratories for the MalE-1 apo measurements. Nearly identical mean FRET efficiencies and standard deviations were determined upon reanalysis (0.49 ± 0.05 versus 0.47 ± 0.06, Supplementary Fig. 8) indicating that user bias was less significant when a global, well-defined analysis procedure was provided over several samples covering a significant fraction of the FRET range (Supplementary Note 2).

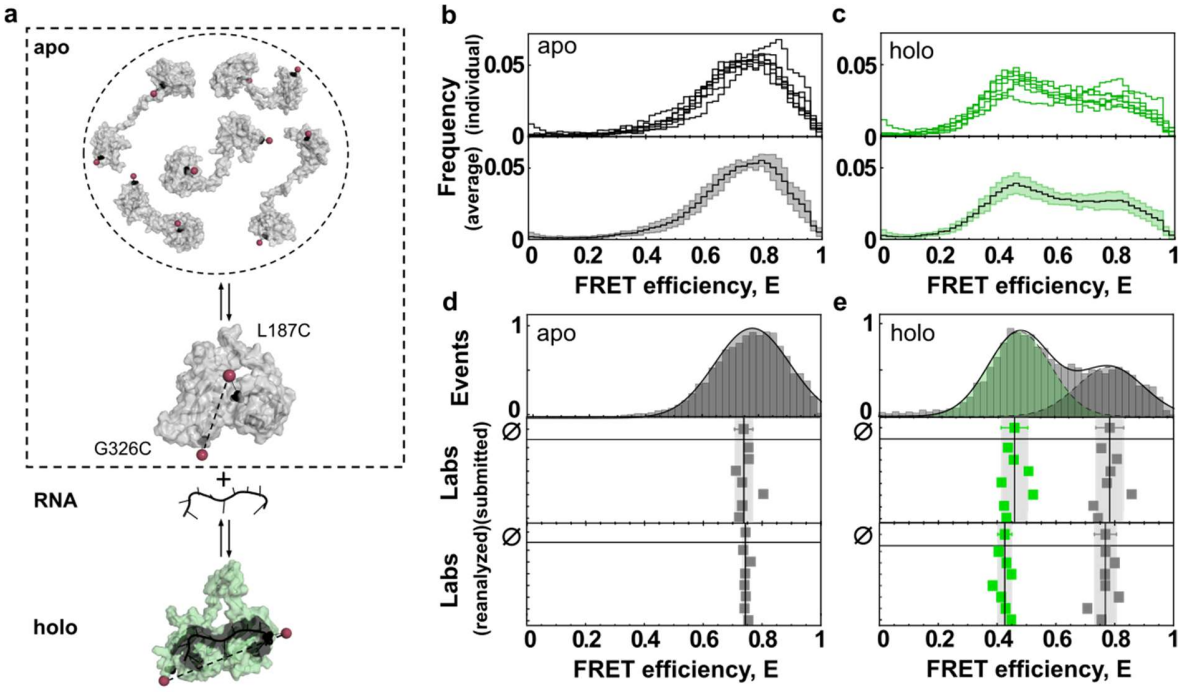


Fig. 2. The experimental system of U2AF2 (RRM1,2) and a comparison of FRET efficiency histograms from 7 different laboratories: (a) Schematic of the dynamics of U2AF2. U2AF2 is comprised of two tandem RNA-binding motifs, RRM1 and RRM2, which are connected by a flexible linker. The apo state (in grey, top) undergoes fast exchange between an ensemble of detached structures of which 5 representative structures are displayed. A slower exchange occurs between the dynamic detached ensemble and a compact conformation (PDB-ID: 2YHO) shown below. The holo state (in green, PDB-ID: 2YH1), shown with a bound U9 RNA ligand, assumes a well-defined, open conformation. Positions used for introducing cysteine mutations for labeling at L187 in RRM1 and G326 in RRM2 are depicted as black spheres with the mean dye position determined by AV calculations indicated by red spheres. **(b-c)** FRET efficiency histograms reported by participating 7 labs for (**b**) apo and (**c**) holo measurements of U2AF2. Top: Individual FRET efficiency histograms and bottom: the average FRET efficiency histogram from the 7 reporting labs (solid line) with standard deviation (light area). **(d)** FRET efficiency *E* histograms of U2AF2 in the apo state. Top: A representative 1-D FRET efficiency histogram of lab 1 fitted to a Gaussian distribution with a mean FRET efficiency of 0.75. Middle: The reported mean FRET efficiencies reported by 7 labs. The mean value from all data sets is 0.739±0.029, shown above with the corresponding standard deviation in grey. Bottom: The extracted mean FRET values after reanalysis of the collected data. After reanalysis, the agreement improved to 0.742±0.008. **(e)** FRET efficiency histogram comparisons of U2AF2 in the holo state. 5 µM of U9 RNA was used to obtain the holo state FRET histogram for U2AF2. Note the decrease in FRET efficiency after binding of RNA to U2AF2. Top: A representative 1-D FRET efficiency histogram of lab 1 fitted to two Gaussian distributions to determine the FRET efficiencies of the different subpopulations, yielding mean FRET efficiencies of 0.44 for RNA-bound and 0.76 for the RNA-free conformation. Middle: The mean FRET efficiencies reported by the 7 labs. The mean values from all 7 the data sets were 0.45±0.04 for the RNA-bound conformation (in green) and 0.78±0.04 for the RNA-free conformation (in grey). Bottom: Reanalysis of the holo measurements yielding values of 0.42 ± 0.02 and 0.77 ± 0.03 for RNA-bound and RNA-free fractions respectively.

For the holo state of U2AF2, good agreement was obtained for the peak positions with a standard deviation of ±0.03 and ±0.02 for the high-FRET and low-FRET peaks, respectively, and only a

minimal improvement resulted from the reanalysis (Supplementary Table 4). In this case, the two populations allowed for a more robust determination of the γ-factor, which can be performed by analyzing the FRET efficiency versus stoichiometry, *S*. In contrast to the good agreement in FRET efficiency, we observed larger variations in the relative amplitudes of the two populations: 0.58±0.08 for the holo state and 0.42±0.08 for the apo population (Fig. 2c, Supplementary Table 4). Such differences are not unexpected due to potentially reduced protein activity, degradation of the ligand, and the high sensitivity of biomolecular dynamics to the experimental conditions, e.g., temperature, ligand concentration, buffer composition, salt concentration or the presence of stabilizers such as BSA (Supplementary Fig. 7).

Characterizing setup-dependent parameters and correction factors. The quality of smFRET experiments is determined by the statistics of the measurement and the performance of the setup to maximize photon collection and thereby minimize shot noise. To this end, we quantified the number of bursts, average count rate, burst duration and number of photons in the donor and FRET channels for the reported MalE measurements from eight labs (Fig. 3a, Supplementary Fig. 9). On average, participants collected 6000 bursts (min: 500, max: 21,000) of molecules carrying both the donor and acceptor fluorophore. The required number of bursts for a smFRET analysis depends on the goal of the experiment. For a simple estimation of an average FRET efficiency from a single population, as performed for MalE, a low number of double-labeled bursts of ~1000 may be sufficient. However, if advanced analysis methods such as time-correlated single photon counting (TCSPC) detection for lifetime analysis, burst-wise fluorescence correlation spectroscopy (FCS) or a photon distribution analysis (PDA) are to be applied to sub-ensembles, higher burst numbers of (>5000) are desired for a robust analysis. Typical count rates per single-molecule event were found to be 60±20 kHz, and an average of 90±40 photons were detected over a typical burst duration of 1.7±0.9 ms (Fig. 3a, Supplementary Fig. 9). The average count rate and burst duration depend on the size of the confocal volume, where smaller sizes typically result in higher count rates but shorter burst durations. Indeed, for the collected data, we observe a negative correlation between the burst duration and the average count rate (Fig. 3b, Pearson's r = -0.58, Supplementary Fig. 10). The large spread of the burst duration arises from the fact that some participants applied a diffraction limited observation volume while others intentionally underfilled the objective lens to create a larger confocal volume with a diameter of ~1 µm (assuming that the labs have adjusted their detection pinhole to correspond with the enlarged excitation volume). We also observed a small positive correlation between the number of detected photons and the burst duration (Fig. 3c, Pearson's r = 0.54, Supplementary Fig. 10). These results indicate that larger confocal volumes, in combination with high irradiances, yield the highest number of photons per burst[38]. Smaller observation volumes generally yield higher count rates and thus shorter inter-photon times, enabling fast transitions on the sub-µs timescale to be resolved[39,40]. Longer burst durations offer the benefit that slower dynamics can be studied.

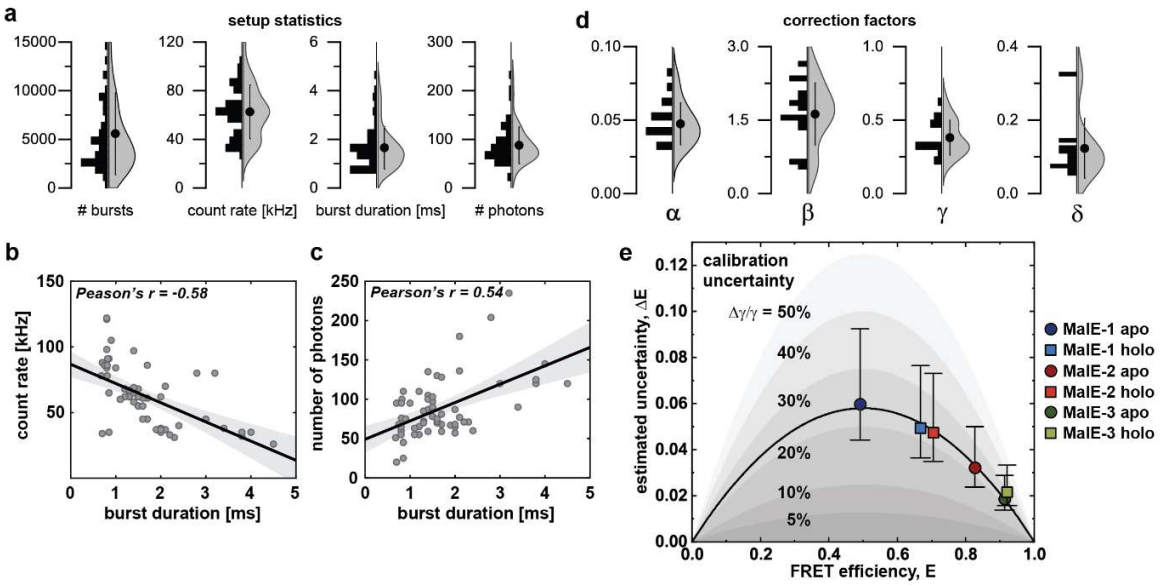


Fig. 3. Setup-dependent parameters and calibration uncertainty. (a) The distribution of the parameters quantifying the statistics of the measurements and the performance of the setups used for both MalE and U2AF2 measurements are shown as histograms and violin plots for the measurements from 8 labs. The circle and whiskers in the violin plot indicate the mean and standard deviation. Sample-dependent distributions of the shown parameters are given in Supplementary Fig. 9. **(b-c)** Pairwise plots of the average count rate **(b)** and number of photons **(c)** against the burst duration. The same datasets are plotted as used for panel **(a).** While the count rate decreases slightly for longer burst durations, a positive correlation is observed for the acquired number of photons per burst and the burst duration, indicating that larger observation volumes result in a higher accumulated signal per molecule. Correlations between all parameters are shown in Supplementary Fig. 10. **(d)** The distributions of the correction factors for the calculation of accurate FRET efficiencies for all the MalE measurements are shown as histograms and violin plots for the measurements from all labs. **(e)** A plot of the standard deviation of the reported FRET efficiencies (as a measure of the experimental uncertainty) against the average FRET efficiency for the MalE mutants 1-3 reveals that lower uncertainties are observed for higher FRET efficiencies. The black line represents a fit of the estimated uncertainties under the assumption that the variations arise solely due to an uncertainty in the γ-factor (see Eq. 1). The inferred relative uncertainty of the γ-factor is ~23%. Shaded areas indicate relative uncertainties of 5-50%. Error bars indicate 95% confidence intervals.

For an accurate analysis of the data, the correction factors for spectral crosstalk (α), excitation flux (β), detection efficiency (γ) and direct excitation (δ) must be determined. Based on data from 16 labs, we plot the distribution of the correction factors used to determine accurate FRET efficiencies for the MalE system in Fig. 3d (Supplementary Table 1). Besides fluorophore properties, the correction factors also depend on setup-specific parameters such as the dichroic mirrors, the emission filters, the detectors, and the excitation wavelengths and power. Nonetheless, we observed a defined distribution for the crosstalk correction factors α of 0.05±0.01, which is mainly determined by the emission filters and type of detectors used for the donor and acceptor detection channels. A larger spread was observed for the correction factor for the excitation flux β of 1.6±0.6 and direct excitation δ of 0.12±0.08. Both factors depend on the ratio of the excitation powers for the donor and acceptor fluorophores. Most participants used about half the laser power for direct

acceptor excitation (45±27 µW) as they used for excitation of the donor fluorophore (78±58 µW) to achieve similar count rates after donor and acceptor excitation. The agreement between the reported FRET efficiency values clearly shows that the various experimental settings are compensated for by the self-consistent correction procedure applied here.

For the detection efficiency correction factor *γ*, we observed an average of 0.4±0.1. The *γ*-factor is arguably the most difficult to determine (see Supplementary Fig. 3b). It depends on the acceptor to donor ratio of the detection efficiencies, g, and the effective fluorescence quantum yields $\phi_F^{\mathrm{eff}} = a_b \phi_F$, where a_b represents the fractions of molecules in the bright state, as $\gamma = g_A\, \phi_{F,A}^{\mathrm{eff}} / g_D\, \phi_{F,D}^{\mathrm{eff}}$ [18]. Like the crosstalk correction term, the detection efficiencies strongly depend on the emission filters and the type of detectors used. Due to the relatively low fluorescence quantum yield of the acceptor, $\phi_{F,A}$ ~ 0.32, compared to that of the donor, $\phi_{F,D}$ ~ 0.72, all labs reported *γ*-factors below 1. Despite the large spread in the different correction factors, we observed very good agreement for the reported FRET efficiencies in our blind study. However, *γ* is also the key factor that limits the consistency between laboratories. This notion is supported by two observations: (i) In Fig. 1d, the spread of FRET efficiency differences, $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$, is smaller (e.g., 0.06 to 0.02 for MalE-1) than for the absolute *E* values in Fig. 1c, suggesting that errors in *E* are systematic rather than random. (ii) The observed spread in reported FRET efficiencies depends on the absolute FRET efficiency measured for MalE (Fig 1c,d). We also calculated the uncertainty due to all parameters in the FRET efficiency calculation using error propagation for cross-talk, direct acceptor excitation, background correction in the donor and acceptor channels. The reported uncertainty can be attributed mainly to the uncertainty in the *γ*-factor (Fig. 3e, Supplementary Note 4). The error of the *γ*-factor, Δ*γ,* propagates into an uncertainty in the reported FRET efficiencies, Δ*E*, as follows:

$$\Delta E = E(1-E)\frac{\Delta\gamma}{\gamma} \qquad (1)$$

Notably, the observed ΔE is well described by Eq. 1 (black line in Fig. 3e), yielding a relative uncertainty of Δγ/γ = 23% (corresponding to an approximate absolute error of Δγ = 0.07). The improved agreement between the measurements upon reanalysis by a single person for U2AF2 (Fig. 2d) indicates that the accuracy of the analysis could be further improved by establishing robust and standardized procedures for the determination of all experimental correction factors, which differ depending on how many populations are present in the measurement and whether the FRET efficiency peak is dynamically averaged (Supplementary Note 2).

Detection and quantification of conformational dynamics in proteins via smFRET. For immobilized molecules, the analysis of dwell-times from the fluorescence trajectories provides access to kinetics on the millisecond to second time-scales (Supplementary Fig. 1)[41–43]. When performing smFRET experiments using confocal detection of freely diffusing molecules, millisecond dynamics can also be measured from a direct analysis of the intensity trajectories (for slowly diffusing molecules)[44,45]. A number of additional approaches can be used for detecting and quantifying faster sub-millisecond conformational dynamics (with the maximum timescale limited by the burst duration) such as FRET-FCS[44,46,47], filtered-FCS[48,49], burst-variance analysis (BVA)[50], FRET-2CDE[51], dynamic PDA[52], FRET efficiency *E* versus fluorescence-weighted average donor lifetime $\langle \tau_{D(A)} \rangle_F$ analysis (*E*-τ plots)[52,53], nanosecond-FCS[54], recurrence analysis of single

particles[55], photon-by-photon maximum likelihood approaches[40,56–59] and Monte Carlo diffusion-enhanced photon inference (MC-DEPI)[60]. To assess how consistently dynamics can be detected, we asked the various groups in this blind study to evaluate whether the protein systems they studied were static or dynamic on the millisecond timescale and which method they used to come to this conclusion (Supplementary Table 6).

The most frequently used methods to evaluate dynamics were the BVA and *E*-τ plots. Both techniques visualize FRET dynamics by comparing the measured data to theoretical expectations for static systems (Fig. 4a/b). BVA detects dynamics by estimating the standard deviation of the FRET efficiency over the time course of the individual bursts, using a predefined photon window (typically around $\gtrsim$ 100 µs depending on the molecular brightness). Due to FRET dynamics, the standard deviation of the FRET signal within a burst (red line in Fig. 4a) is higher than expected from shot noise (black line in Fig. 4a), which becomes visible as a deviation (apparent dynamic shift, ds) from the shot-noise limited standard deviation of the apparent FRET efficiency (which is a semi-circle in shape)[50].

In the *E*-τ plots, the observed FRET efficiency determined via intensity (the *y*-axis in Fig. 4b) is a species-weighted average and, in the presence of dynamics, the position along this axis depends on the fraction of time spent in the respective states. The fluorescence lifetime of the donor ($\langle\tau_{D(A)}\rangle_F$, the *x*-axis in Fig. 4b) is a photon-weighted average because only a single lifetime can be determined from the single-molecule lifetime data. It is weighted towards the lifetime of the lower FRET state as the majority of photons are emitted from the donor in the low FRET efficiency state[52,53], shifting the data to the right of the 'static' FRET-line. *E*-τ plots can detect dynamics on the ns to ms timescale. Note that, for the experimental data, we have included an additional correction to the 'static' FRET-line that accounts for the distance fluctuations due to the flexible dye linkers of 6 Å, resulting in a slightly curved line[61]. To quantify the dynamics between two distinct states, a theoretical 'dynamic' FRET-line (red line in Fig. 4b) can be calculated and overlaid on the plots. Again, the apparent dynamic shift, ds, is defined as the deviation of the observed data from the theoretical static line (Fig. 4b). See Supplementary Note 6 for details.

We previously showed that MalE exhibits slow ligand-driven dynamics on the sub-second timescale between high- and low-FRET states (Supplementary Fig. 1)[62]. Here, we investigated whether the apo and holo states of MalE are undergoing dynamics faster than or on the timescale of the diffusion time of 1-3 ms. Both techniques reveal that the conformations of MalE exhibit no large FRET-fluctuations on the ms timescale (Fig. 4c,d; Supplementary Fig. 11). Almost all groups confirmed this assessment for all MalE samples (Supplementary Table 6). Three groups concluded that MalE is dynamic without further justification. To investigate potential dynamics in more detail, we determined the dynamic shifts for a subset of the data (eight labs for BVA, Fig. 4e and five for *E*-τ, Fig. 4f, Supplementary Note 5, Supplementary Table 7). As a static control, we determined the dynamic shift of the dsDNA rulers used in Hellenkamp *et al.*[18] (mean ± one standard deviation as determined from Labs 1 and 2) shown in grey in Fig. 4e,f (Supplementary Table 8). Interestingly, no apparent dynamic shift exceeding the dsDNA reference was observed when using BVA for all MalE mutants. From the *E*-τ plots, however, there is an apparent dynamic shift, especially for MalE-1 of ~0.05, that clearly exceeds what would be expected for a static system or even what is predicted for potential dynamics between the apo and holo conformation (Fig. 4f, red lines, Supplementary Note 6). Hence, some labs categorized MalE as dynamic. The

origin of this apparent shift, which must originate from dynamics that are faster than ~100 µs, will be discussed in detail below.

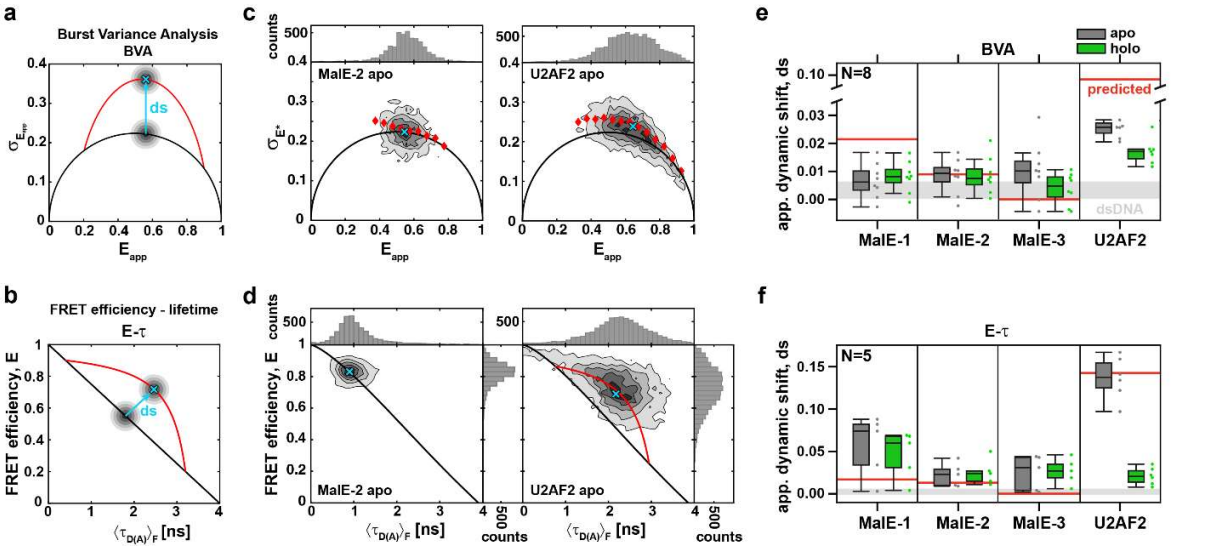


Fig. 4. Detection and characterization of conformational dynamics on the sub-millisecond timescale in MalE and U2AF2: (a-b) Schematic representations of burst variance analysis (BVA) and *E*-τ plot. (a) In BVA, the standard deviation $\sigma_{E_{app}}$ of the apparent FRET efficiency E_{app} is compared to the shot-noise limit. Single-molecule events with conformational dynamics show increased variance and follow the dynamic line (red). The dynamic shift ds is defined as the excess standard deviation compared to the static line. (b) In the *E*-τ plot, the intensity-based FRET efficiency *E* is plotted against the intensity-weighted average donor fluorescence lifetime, $\langle\tau_{D(A)}\rangle_F$. Molecules undergoing dynamics are shifted from the static line (black) and follow a dynamic FRET-line (red, see text)[61]. For a given population, the dynamic shift is defined as the displacement of the population orthogonal to the static FRET-line. See Supplementary Note 5 for details. **(c)** BVA of MalE-2 labeled with Alexa546-Alexa647 without maltose (apo, left) and U2AF2 labeled with Atto532-Atto643 without RNA (apo, right). Here, the BVA is based on a photon binning of 5 photons. Red diamonds indicate the average standard deviation of all bursts within a FRET efficiency range of 0.05. The mean positions of the populations (cyan crosses) were determined by fitting a two-dimensional Gaussian distribution to the data (Supplementary Note 5). **(d)** The plots of the FRET efficiency *E* versus intensity-weighted average donor lifetime $\langle\tau_{D(A)}\rangle_F$ of the same measurement as in **(c)**. The static lines are slightly curved as they account for the flexibility of the dye linkers[52]. The donor-only population was excluded from the plot. For MalE-2, the population falls on the static FRET-line, while a clear dynamic shift is observed for U2AF2. The end points of the dynamic FRET-line for U2AF2 were determined from a sub-ensemble analysis of the fluorescence lifetime decay. **(e-f)** The dynamic shift of the peak of the population was determined graphically from BVA (8 labs for MalE, and 7 labs for U2AF2 respectively) and *E*-τ (5 labs) plots (see Online Methods). For U2AF2 in the holo state, the dynamic shift was assessed only for the low-FRET RNA-bound population. All labs consistently detected the highest dynamic shift for U2AF2 in the apo state. A significant dynamic shift was also consistently detected for MalE-1. Boxes indicate the median and 25%/75% quartiles of the data. Whiskers extend to the lowest or highest data point within 1.5-times the interquartile range. The grey area indicates the dynamic shift obtained for the double-stranded DNA used in a previous benchmark study[18] based on measurements performed in lab 1 for BVA (ds_{DNA} = 0.0033 ± 0.0033) and lab 2 for the *E*-τ plot (ds_{DNA} = 0.0026 ± 0.0044). The horizontal red lines indicate the expected dynamic shift for a potential

conformational exchange between apo and holo states. We computed the expected change of FRET efficiency using their structural models in the PDB (Supplementary Note 6 and Supplementary Table 9).

In contrast to MalE, all groups found U2AF2 to be dynamic as was expected for two domains connected by a flexible linker (Fig. 4c-f, Supplementary Table 6). The ligand-free apo state shows pronounced deviations from the behavior for static molecules both in the BVA and *E*-τ plots, while the RNA-bound holo state shows a significant apparent dynamic shift for BVA but not for the *E*-τ analysis (Fig. 4c-f). Due to the existence of a significant fraction of apo-protein and the overlap between the apo and holo populations, it was challenging to assess whether the holo state is truly static or dynamic, although a clear apparent shift is observed. In summary, even though U2AF2 is a very challenging test case, dynamics were unambiguously detected in all labs demonstrating the reliability of smFRET for investigating dynamic systems.

Accuracy of FRET-derived distances in proteins with respect to structural models. After determining FRET efficiencies of different conformations in a protein, the next step is (often) to convert these FRET efficiencies into distances and compare them to what is expected from structures or to use them as distance constraints in integrative FRET-assisted structural modeling[1,5,7,15,63]. The smFRET experiments yield the FRET efficiency as a result of dynamically, non-linearly averaged distances due to the flexible linker used to attached the fluorophore to the molecule. Fast and robust ways of accurately modeling the fluorophore positions and volumes accessible to the fluorophore attached to the biomacromolecule is a topic of ongoing investigation[5,63,64]. To assess the accuracy of our measurements, we measured the fluorescence lifetimes and time-resolved anisotropies for each labeling position and applied the accessible volumes (AV) approach[5,6,65,66] that employs a coarse-grained dye model to estimate the FRET efficiency averaged model distance $R_{\langle E\rangle}^{\mathrm{model}}$ between the two dyes. For this, all possible positions of the donor and acceptor fluorophores are averaged, taking into account distinct linker conformations and steric hindrances of the protein (Fig. 5a-c; see Online Methods and reference 6 for details). For the AV calculations, we assume fast rotational averaging and slow position averaging on the timescale of the fluorescence lifetime. We note that prediction of the distance measured via FRET values based on the flexibility and attachment point of the fluorophore is an area of active research and alternative methods are being developed such as using rotamer libraries[67] or molecular-dynamics-simulations[68,69].

The experimental FRET efficiencies $\langle E\rangle$ for MalE from all labs (Fig. 2) were utilized to determine $R_{\langle E\rangle}$ for each lab (Table 1 and Supplementary Table 3) using the Förster equation (Eq. 2):

$$R_{\langle E\rangle} = R_0 \left(\frac{1}{\langle E\rangle} - 1\right)^{\frac{1}{6}} \qquad (2)$$

The Förster radius of Alexa546-Alexa647 on MalE was determined to be R_0 = 6.5±0.3 nm (see

Supplementary Note 7). In Fig. 5d, we display the correlation between the experimental observable $R_{\langle E \rangle}$ and predicted $R_{\langle E \rangle}^{\mathrm{model}}$ using the known, well-defined structures of the apo and holo states showing an uncertainty of 3-5 Å over all mutants. In agreement with the predictions by Peulen et al.[70], this accuracy is achieved despite stochastic protein labeling, which could result in drastically different charge environments and accessible volumes of the fluorophores depending on their locations on the protein, as evidenced by the varying dye behavior at different labeling positions (Fig. 5). Over the course of this study, three labs studied two additional MalE mutants (MalE-4: K34C-N205C and MalE-5: T36C-N205C) which were designed to provide a larger FRET efficiency contrast between the apo and holo states, complementing the results of the other variants (Table 1).

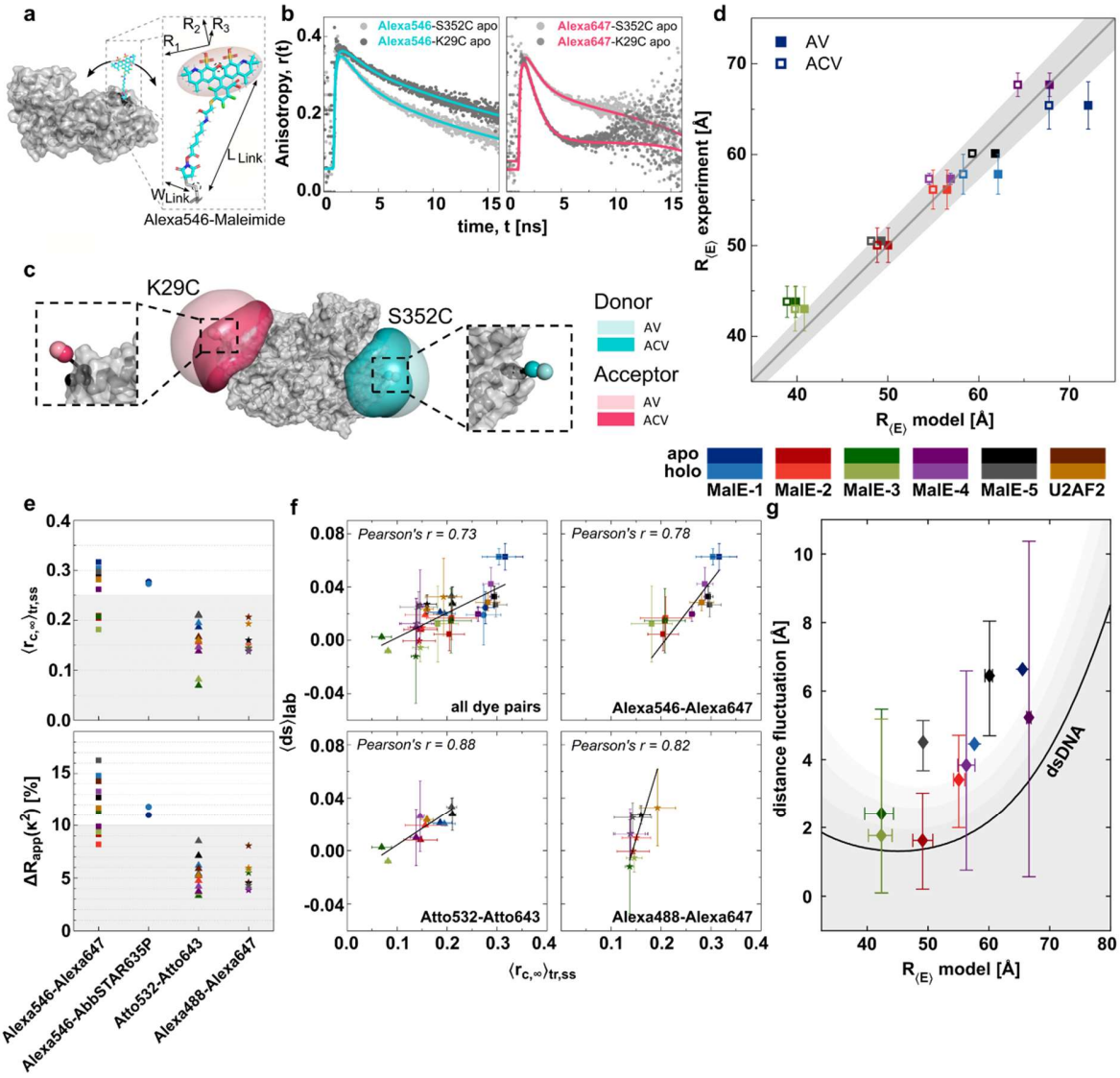


Fig. 5. Assessing the accuracy of FRET-derived distances in MalE. (a-d) Accessible volume calculations and model-based interdye distances: **(a)** Scheme of the Alexa546 fluorophore attached to MalE (PDB 1OMP) showing the parameters used for the accessible volume calculations. The fluorophore is modeled as an ellipsoid with a flexible linker using the AV3 model[6] (Supplementary Table 10). **(b)**

Fluorescence anisotropy decays of the donor fluorophore (Alexa546, left) and acceptor fluorophore (Alexa647, right) at the two labeling positions K29C and S352C on mutant MalE-1. The anisotropy decays were obtained from single-cysteine mutants labeled only with the donor or acceptor fluorophore respectively. Strong hindrance of the rotation due to sticking to the protein surface, as indicated from the high vertical offset and slow decay of the anisotropy, are detected for the Alexa546 at both positions and Alexa647 at position S352C. Solid lines represent fits to a model of two or three rotational components (Supplementary Tables 11 and 12). The analysis was performed as described in Supplementary Note 8. **(c)** Accessible volumes for Alexa546 (cyan) and Alexa647 (pink) at labeling positions 352 and 29 calculated using the AV model as described in panel (a). The contact volume close to the biomolecular surface (shown as a darker shade) is used in the ACV model by weighting the occupancy of the contact volume based on the residual anisotropy. The zoom-ins show the mean positions of the dyes based on the AV (light shade) and ACV (darker shade) model. In the ACV model, the position of the dyes is biased towards the protein surface, resulting in a reduction of the interdye distance for the given labeling positions. See Online Methods for details. **(d)** Comparison of the experimentally obtained FRET-averaged distance $R_{\langle E \rangle}$ with the theoretical model distances using the AV (filled squares) and ACV (empty squares) calculations. Errors of the experimental distances represent the standard deviation over all labs. The solid line represents a 1:1 relation and the grey area indicates an uncertainty of ± 3 Å for a Förster radius of R_0 = 6.5 nm. The two additional mutants, MalE-4 and 5, labeled at positions K34C-N205C and T36C-N205C, were measured by two labs. The agreement between the model and experiment (determined using the average root-mean-square deviation) decreases from 3 Å for the AV model to 2 Å for the ACV model. **(e)** Detection of dye-specific protein interactions. (Top) The five MalE mutants and U2AF2 were labeled with four different dye combinations (Alexa546-Alexa647, Alexa546-AbbSTAR635P, Atto532-Atto643 and Alexa488-Alexa647) and measured by three different labs to determine the donor-acceptor-combined residual anisotropy from time-resolved (tr) and steady state (ss) anisotropy measurements, $\langle r_{c,\infty} \rangle_{tr,ss}$. (Bottom) From the residual anisotropy of the donor and acceptor fluorophores, the distance uncertainty relating to the orientation factor κ^2, $\Delta R_{app}(\kappa^2)$, was estimated for the different dye pairs, as described in the Supplementary Note 8. Based on the distance uncertainty, $\Delta R_{app}(\kappa^2)$, a threshold is derived to filter the datasets based on the sticking propensity of the dyes, a maximum allowed distance uncertainty of $\leq 10\%$ (shaded grey region) leads to a dye-independent threshold for $\langle r_{c,\infty} \rangle$ of 0.25. **(f)** The apparent dynamic shift $\langle ds \rangle$ shows a strong correlation with the combined residual anisotropy $\langle r_{c,\infty} \rangle$ over all measured dye pairs (top left, Pearson correlation coefficient 0.73), indicating that sticking interactions can lead to a false-positive detection of conformational dynamics. A higher correlation between $\langle ds \rangle$ and $\langle r_{c,\infty} \rangle$ is observed when the dyes pairs are analyzed separately. **(g)** The structural flexibility of MalE was estimated based on the residual dynamic shift after filtering using the distance uncertainty threshold shown in **e**. The residual dynamic shift is converted into a corresponding distance fluctuation assuming a two-state dynamic exchange that is symmetric around the center distance $R_{\langle E \rangle}$ (see methods). The residual distance fluctuations obtained for control measurements on double-stranded DNA (ds_{DNA} = 0.0026 ± 0.0044) is shown as a black line (gray areas represent the 1σ, 2σ and 3σ confidence intervals).

Closer inspection of Fig. 5d reveals the largest deviation for MalE-1, which also showed a significant dynamic shift in the *E*-τ plot (Fig. 4f and Supplementary Fig. 11). Therefore, we investigated whether dye-protein interactions play a role for the donor or acceptor dye by measuring the fluorescence lifetime and the time-resolved and steady state anisotropies of single-cysteine variants of MalE (Supplementary Note 8, Supplementary Table 5, 11 and 12, and Fig. 5b). These results show that labeling at residue 352 strongly promotes sticking to the protein surface indicated by multiexponential fluorescence lifetimes and a high residual anisotropy, r_∞, for both the donor and acceptor fluorophores (r_∞ > 0.25), while labeling at residue 29 only shows sticking for the donor ($r_{\infty,D}$ > 0.30, $r_{\infty,A}$ ~ 0.12). However, at other positions (e.g., residue 186), free rotation is possible for both dyes (Supplementary Tables 5 and 11). These position-specific

interactions can cause the deviations of the experimentally determined distances from the AV model (Fig. 5d, Table 1) and the apparent dynamic shift for mutant MalE-1 (Fig. 4f). A more accurate prediction of the model distances is obtained when the dye sticking is accounted for using the accessible contact volume, ACV[63], approach (Fig. 5c). When labeling the protein on opposite sides, the dye-surface interactions in the ACV model generally results in a reduced model distance (Fig. 5d, Table 1), which leads to a significant improvement of the accuracy for the outlying mutants.

It has been previously suggested to use the combined residual anisotropy of D and A, computed via $r_{c,\infty} = \sqrt{r_{\infty,D} \cdot r_{\infty,A}}$, as a criterion for filtering out dye-related artifacts in FRET-assisted structural modeling with an empirical threshold of $r_{c,\infty} < 0.2$[13,71]. To further investigate dye-specific sticking, three labs performed measurements of different MalE mutants with additional dye pairs (Alexa546-AbbSTAR635P, Atto532-Atto643 and Alexa Fluor 488 (Alexa488)-Alexa647) and determined the residual anisotropies and distance uncertainties based on the orientation factor κ^2 (Fig. 5e, top; Supplementary Table 13 and 14, Supplementary Note 8 and 9). These results depend on the dye-pair, protein and labeling position investigated and have to be addressed individually for the specific system being studied. In this case, the dye pair Alexa546-Alexa647 showed the highest combined anisotropies (Supplementary Fig. 12a, Supplementary Table 13). This is attributed mainly to the donor dye Alexa546 as the combined anisotropy remains high for a different acceptor (Alexa546-AbbSTAR635P) but is reduced markedly for another donor fluorophore (Alexa488-Alexa647). To derive a robust and well-defined threshold for recognizing measurements with dye artifacts, we determined the uncertainty in the FRET-derived distances, $\Delta R_{app}(\kappa^2)$, that originates from the uncertainty of the orientation factor κ^2. Previous approaches have estimated the uncertainty in κ^2 from the residual anisotropy in terms of rotational restrictions (wobbling-in-a-cone model)[71–74]. Here, we used a 'diffusion with traps' model suggested by S. Kalinin, which assumes two dye populations (free and trapped) and relates the residual anisotropies to the fraction of dyes sticking to the surface of the labeled biomolecule (for details, see Supplementary Note 9). Based on the estimated distance uncertainty, we propose a threshold of $\Delta R_{app}(\kappa^2) < 10\%$ to identify measurements with dye-related artifacts (Fig. 5e, bottom). This threshold corresponds to a combined residual anisotropy of 0.25, similar to the previous empirical threshold value of ~0.2[13,71].

Next, we investigated whether dye sticking could indeed cause an apparent dynamic shift in the E-τ plot as seen for MalE-1 with the dye pair Alexa546-Alexa647 (Fig. 4f). For the effect to be observable in the E-τ plot, the exchange between the free and trapped species must occur faster than the diffusion time of ~1 ms, otherwise the two species would be observable as individual peaks. We observed a correlation between the lab-averaged apparent dynamic shift $\langle ds \rangle$ and combined residual anisotropy $\langle r_{c,\infty} \rangle$ over all dye pairs (Pearson's r = 0.73), with a stronger correlation being observed for each dye pair individually (Fig. 5f). As conformational dynamics should be independent of the labels used, we conclude that dye sticking is responsible for the observed apparent dynamic shift. Interestingly, the x-intercept of the linear fit is between 0.1 and 0.2, suggesting a dye-dependent anisotropy threshold needs to be considered. When applying the criteria $\langle r_{c,\infty} \rangle < 0.25$ to MalE-1 (Supplementary Fig. 12b), only the dye pair Atto532-Atto643 could be used, which also showed a significantly reduced apparent dynamic shift (Supplementary Fig. 12c). A lifetime analysis of Alexa546 donor only molecules from MalE-1 showed donor quenching that is not observed at other positions, which confirms that labeling at position 352 is

especially problematic (see Supplementary Fig. 12c, Supplementary Note 10 and Supplementary Table 5).

Using the above criteria of $\langle r_{c,\infty} \rangle < 0.25$ to minimize the influence of dye artifacts on the dynamic shift, we were interested in finding out whether the observed dynamic shift for the other MalE mutants could be indicative of low-amplitude, fast conformational fluctuations. A *p*-test analysis between the dynamic shift for DNA rulers and protein samples ($p < 0.05$) indicated that the dynamic shift calculated after filtering out dye artifacts is still significant for various protein variants (Supplementary Note 11, Supplementary Table 8). To estimate the magnitude of the conformational fluctuations necessary to generate the observed dynamic shifts (Fig. 4f and Supplementary Table 8), we assume that the dynamics occur between two nearby states with interdye distances of $R_{\langle E \rangle} \pm \delta R$, where δR is the amplitude of the fluctuation[61] (Fig. 5g, Supplementary Note 12 and Supplementary Table 8). This inferred distance fluctuation must be interpreted as an upper bound for the conformational flexibility because other factors are likely to contribute to the dynamic shift such as calibration errors, dye blinking or photoisomerization. To account for experimental errors that induce false-positive dynamic shifts, we consider the dynamic shift obtained from dsDNA molecules as the lower limit (black line in Fig. 5g, ds_{DNA} = 0.0026±0.0044, see Supplementary Note 12), which defines the current detection limit for dynamics in smFRET experiments. The MalE variants 1, 4 and 5 clearly exceed the dynamic shift observed for dsDNA by 2-3 Å (Fig. 5g, Supplementary Fig. 13, Supplementary Table 8). Consistent with the smFRET results, all-atom MD simulations of MalE using the ff14SB force field[75], which is widely used for folded proteins, clearly indicate the presence of small thermally-induced conformational fluctuations with a standard deviation of up to ~3 Å at the labeling locations used for MalE-1, MalE-4 and MalE-5. This is larger than the typical fluctuations on the order of 1 Å[76] and leads to a broadening of the inter-residue distance distributions for the used FRET pairs (Supplementary Note 13). We conclude that the residual dynamic shift observed in the experiments can be sufficiently explained by a combination of measurement uncertainty and small-scale structural fluctuations. Note that such small-scale distance fluctuations can be amplified in FRET experiments because the dye linker can act as a lever arm, leading to an enhancement in the FRET contrast if labeling positions are chosen appropriately. A detailed discussion of the theoretical limits for detecting dynamics in smFRET experiments using BVA or the E-τ is given in Supplementary Note 14.

Quantitative analysis of U2AF2: The structural characterization of conformationally flexible U2AF2 is much more complex and a simple distance comparison as for MalE is not possible. Nonetheless, we asked ourselves what information smFRET measurements could provide for such a dynamic system. We first surveyed the structural information available on the conformational ensemble of apo U2AF2 determined using NMR and SAXS[29]. The high flexibility of the linker allows for a heterogeneous ensemble of possible conformations (Fig. 6a). To assess how this conformational heterogeneity translates into the expected smFRET distributions, we quantified the FRET efficiency for each of the 200 conformers available from the NMR/SAXS derived full ensemble of apo U2AF2[29] using AV calculations. Notably, conformations with similar center-of-mass (COM) distances between the domains could show vastly different FRET efficiencies (Fig. 6a-b). This occurs because rotations of the domains can result in the dyes

pointing towards or away from each other (Fig. 6a, right). Due to this degeneracy, the single-distance information provided here is insufficient to capture the full structural complexity of the apo state.

As expected, the significant dynamic shift observed in the smFRET experiments clearly supports the presence of conformational dynamics in the apo state (Fig. 4d-f). To decipher the different kinetics involved and their temporal hierarchy, we applied three analyses that are sensitive to dynamics on different timescales. First, the full conformational heterogeneity of U2AF2 was investigated using the donor fluorescence decay. We infer the distribution of interdye distances for the apo and holo states using a model-free maximum entropy method (MEM, Supplementary Note 15)[77]. To test the consistency between the distance distributions provided by the FRET lifetime analysis and by the NMR/SAXS data, we used the NMR/SAXS full structural ensemble as the prior distribution for the MEM (Fig. 6c-d and Supplementary Note 15). This analysis yielded consistent results for the three dye pairs studied for U2AF2. Notably, the MEM analysis revealed peaks in the probability density at the expected distances for the compact apo conformation and RNA-bound holo structure for all dye combinations (Fig. 6d, dashed lines and Supplementary Note 15). We note that the fluorescence lifetime analysis resolves states on the ns time scale and is therefore less sensitive to dynamic averaging.

Secondly, to assess the dynamics on the microsecond timescale, three groups performed a FRET-FCS or filtered-FCS analysis for the Atto532-Atto643 labeled protein and found at least two relaxation times of 9±3 µs and 300±90 µs (Fig. 6e, Supplementary Table 15, Supplementary Note 16). Control experiments using different dye combinations revealed consistent dynamic timescales (Supplementary Fig. 15). We assign these processes to the fast dynamics of the detached domains and the slower interconversion between compact conformations within the conformational ensemble.

Lastly, we investigated dynamics on the millisecond timescale using a dynamic photon distribution analysis (PDA). Here, we performed a global analysis of the apo and holo measurements using the kinetic model shown in Fig. 6c (Supplementary Note 17 and Supplementary Table 16). We treat the apo state as a two-state system with slow dynamics between a detached ensemble and a well-defined, compact apo conformation. The rapid dynamics within the detached ensemble is empirically described using a broad, static distribution. For the holo measurement, we account for the residual population of apo molecules. Exchange between the holo and apo state is not relevant as the binding and dissociation of the RNA occurs on timescales >100 ms[36]. This global model including all known information provides an approximate description of the measured FRET efficiency histograms (Fig. 6f, Supplementary Fig. 16). The dynamic PDA analysis returned a relaxation time of ~10 ms for the dynamics between the detached ensemble and compact apo state with a very small amplitude (orange curve, Supplementary Fig. 16, Supplementary Table 16). We were also able to accurately determine an interdye distance of $R_{\langle E \rangle}$ = 61 Å in the RNA-bound holo state, which is in very good agreement with the RNA-bound conformation of 63 Å (PDB: 2YH1).

Discussion

The presented results of our blind study clearly demonstrate that smFRET can consistently provide accurate distances of conformational states and reliable information on dynamics in proteins. All experiments were performed using established experimental procedures and analyzed with freely-available data analysis routines[5,6,34,35,78–80], indicating that the presented experiments and the conclusions drawn are accessible to groups with similar technical expertise. Despite the challenges of dealing with proteins samples, we could achieve a similar precision in measured FRET efficiencies for both systems over a large part of the dynamic range of FRET as reported previously for stable oligonucleotide structures[18] (between ± 0.02 and ± 0.06) (Table 1). The high level of consistency for qualitative detection of large-scale sub-millisecond dynamics in U2AF2 and exclusion thereof for MalE shows that the community is well positioned to deal with dynamic protein systems. In addition, we could establish the wide range of timescales and hierarchy of the exchange dynamics observed in U2AF2. The investigation of the complex dynamics could be improved by using multiple labeling positions to measure additional intramolecular distances[5,13,17,81–83]. Consistent results regarding the dynamic timescales were provided by different laboratories using a correlation analysis, and further improvements would be expected when the experimental conditions are better controlled (Supplementary Fig. 7 and Supplementary Table 15).

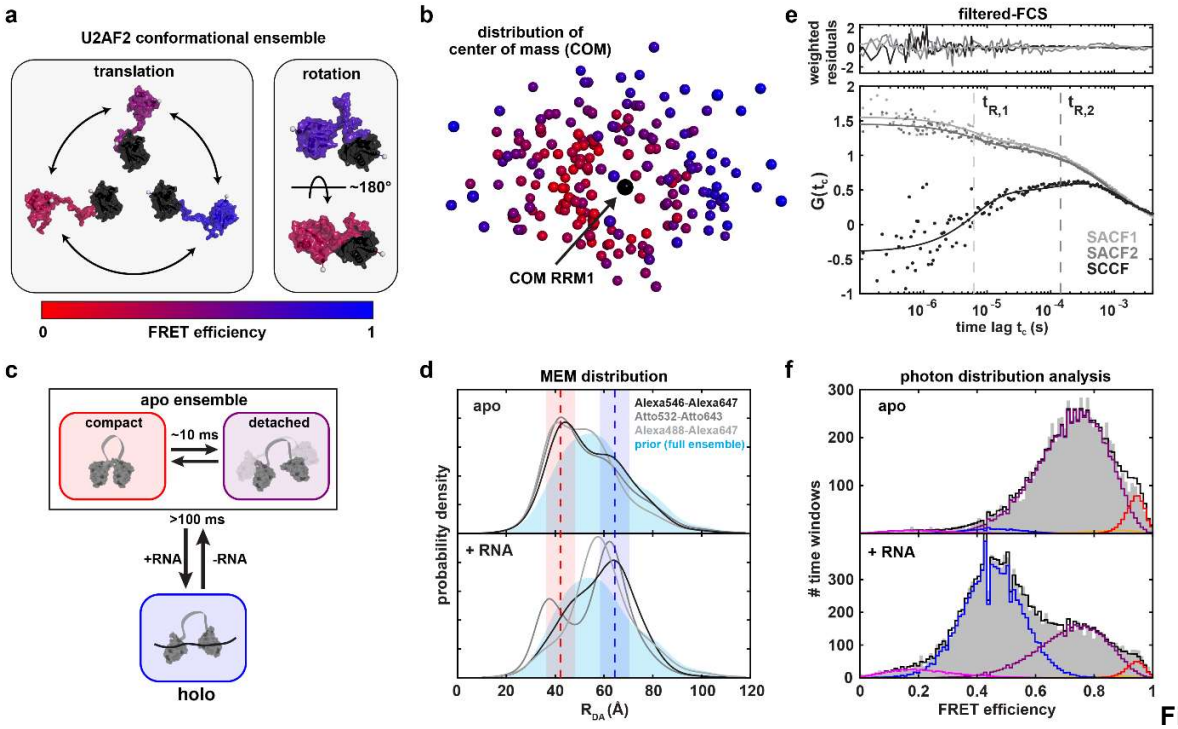


Fig. 6. Structural characterization of U2AF2 labeled with Atto532-Atto643 at positions 187 and 326. (a) Structural flexibility in the conformational ensemble of U2AF2 is given by translational (left) and rotational (right) movement of the two domains. Representative structures are taken from the ensemble determined using NMR and SAXS measurements by Huang et al.[29]. **(b)** Degeneracy of structural states in FRET measurements. The position of the two domains of U2AF2 is illustrated by the center-of-masses of the C_α

atoms (COM) in RRM2 (residues 150-227, colored) with respect to RRM1 (residues 260-329, black) for the 200 structures of the conformational ensemble from Huang et al.[29]. The COM of RRM2 is color-coded according to the FRET efficiency of the conformation based on AV3 calculations. Structures with similar COM distances can exhibit different FRET efficiencies due to rotation of the domains. **(c)** A schematic of the kinetic model used for the global dynamic PDA of U2AF2. In the apo state, the protein fluctuates between two states, a defined compact state and the detached ensemble. The rapid dynamics within the detached ensemble are not seen in PDA due to kinetic averaging and the degeneracy of the ensemble with respect to the FRET efficiency (see panel b). The holo state is populated by binding of RNA. Exchange between the apo and holo states occurs on timescales >100 ms (estimated using the known K_d) and is thus too slow to be visible in the diffusion-based smFRET experiments. **(d)** Distance distributions obtained from a donor fluorescence decay analysis by a model-free MEM approach (Supplementary Note 15). The distance distribution from the full NMR/SAXS ensemble[29] (shown in light blue) was used as the prior distribution. The expected interdye distances for the resolved structure of the compact apo and open holo states are shown as red and blue dashed lines (PDB: 2YH0, 2YH1) with the shaded areas indicating the distance broadening due to the flexible dye linkers of 6 Å. The distribution in the donor-acceptor distance R_{DA} obtained by the MEM analysis for different dye pairs (see grey shading) is shown. See Supplementary Note 15 for details. **(e)** Filtered fluorescence correlation spectroscopy reveals conformational dynamics in the U2AF2 apo ensemble on two timescales, $t_{R,1}$ = 9±3 µs and $t_{R,2}$ = 300±90 µs, averaged over all reporting labs (results from lab 1 are shown). The two species were defined at the lower and upper edge of the FRET efficiency histogram shown in Fig. 2b, top panel, by selecting bursts with $E \leq 0.6$ and $E \geq 0.9$ respectively. The two species-autocorrelation functions (SACF) and the two (one correlated with two and two correlated with one) species-cross-correlation functions (SCCF) were globally fit to a single-component diffusion model with two kinetic relaxation times (see methods and Supplementary Note 16 for details). For clarity, only one of the two SCCFs is shown. The weighted residuals are shown above. **(f)** The global PDA analysis was performed globally over both apo (top) and holo (bottom) measurements using time windows of 0.5, 1, 1.5 and 2 ms (the displayed histograms correspond to a 1 ms time window), resulting in a global reduced χ^2 of 1.69. A relaxation time of ~10 ms for the dynamics between the detached ensemble and compact apo state with a small amplitude was determined (orange curve). See Supplementary Fig. 16 for all histograms and Supplementary Note 17 for details of the analysis.

The high level of consistency is especially notable given the diversity of the setups (Fig. 3 and Supplementary Fig. 2) and the number of difficulties and pitfalls that can occur. The largest contribution to the spread in the reported mean FRET efficiencies was caused by differences in data analysis that can introduce systematic errors. This is demonstrated by investigating the FRET efficiency changes ($\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$) instead of the absolute FRET efficiency values (Fig. 1d), which reduced the spread by a factor of ~ 3. Having a single person reanalyze the data lead to a similar decrease in the uncertainty of the FRET efficiency for the apo state of U2AF2 (Fig. 2d). The most commonly used calibration procedures are γ determination according to Lee *et al.*[34], using a linear regression of $1/S_{app}$ versus E_{app}, and γ determination according to Kudryavtsev *et al.*[35] via *E-τ* calibration (Supplementary Note 2). In the first approach, multiple samples are needed for the calibration and either requires uniform fluorophore properties across all samples or individual corrections made to the samples that deviate. In the second approach, the system needs to be static. A generalized protocol with unambiguous instructions for each of the calibration steps and minimized number of user-dependent steps would alleviate calibration related issues and further enhance the accuracy of FRET measurements. However, the optimal approach depends on the properties of the measured system, making determination of a generalized protocol challenging.

Upon determination of an accurate FRET efficiency, the next step is to convert FRET-efficiency values to inter-dye and inter-base distances as discussed previously in reference [18]. Using the structural model for MalE, we obtained reproducible distances with a precision of 3 Å and an accuracy of 5 Å against structural models (Table 1), values similar to what was determined for dsDNA samples. This is a very positive outcome, given that the DNA standards featured a consistent, homogenous chemical environment for the DNA labeling positions, which was in strong contrast to the much more variable dye environment experienced in the measured proteins.

To improve the distance determination further, two important factors were shown to be useful. First, the interaction of the fluorophores with the surface needs to be included in the accessible volume calculations (Fig 5d)[63]. Secondly, only dye-pairs with a combined residual anisotropy of $r_{c,\infty} < 0.25$ should be used (Fig 5f). By comparing measurements on several dye-pairs, we now give experimental support for the value of $r_{c,\infty} \lesssim 0.25$, in line with previously given empirical thresholds[13,71]. In addition, proteins often exist within a family of conformations and thus a distribution of distances is necessary to properly describe the system. This can be seen, for example, by the lifetime distribution of U2AF2, where a MEM was used to estimate the conformational distribution of the ensemble (Fig. 6d). Determining how to best deal with distance distributions for conformational ensembles is one of the challenges for structural biology.

Investigating different dye pairs allowed us to select samples where dye artifacts are minimized, thereby leading to more accurate and robust FRET efficiencies as well as the reliable detection and quantification of the dynamics. A careful inspection of the *E*-τ plots for MalE raised the question whether a significant deviation from the static FRET-line was observable, implying the existence of FRET dynamics. Therefore, we investigated the detection limits for conformational dynamics using FRET with a subset of laboratories. We note that dynamic FRET shifts can have several origins and may not be due to conformational motions of the protein. For example, linkers used to attach the fluorophore to the molecule or structural instabilities induced by the labeling, as shown by Sánchez-Rico *et al*. for U2AF2, can lead to FRET dynamics[37]. Additional influences come from donor and acceptor quenching, acceptor blinking and photobleaching, and dye sticking (as shown in Fig. 5f)[44]. Thus, in some cases, it can be advisable to verify the key findings in smFRET measurements with at least two dye pairs and/or with different residue combinations in the protein. Contributions to the dynamics shift that have origins other than FRET dynamics can have both positive as well as negative influences on the apparent dynamics shift values. There, we verified with dsDNA structures, which we treat as relatively static biomolecules, that the observed average shift was small (~0.003 in both the BVA and *E*-τ plots). Once the non-FRET-dynamic contributions could be minimized, we still observed significant residual dynamic shifts for MalE. Interpreting these shifts as coming from small-scale conformational dynamics, we establish here a current lower limit for the detection of structural changes via smFRET on the order of ≤ 5 Å. Fluctuations having a similar magnitude were also observed at the labeling positions of various MalE variants in an all-atom MD simulation (Supplementary Note 13).

Conclusions

The consensus of the smFRET data from 19 laboratories on two protein systems exhibiting dynamic behavior on different timescales offers strong support for the use of smFRET as a robust, versatile and quantitative tool for protein distances and dynamics. Deviations in FRET efficiency measured by the various groups were similar to what was determined using DNA standards. One factor that could improve the consistency between laboratories would be a more robust determination of the detection-correction factors required for calculation of setup-independent accurate FRET efficiencies. We also demonstrated that smFRET allows one to detect and characterize conformational dynamics in proteins and can disentangle the latter from dye quenching, blinking, photobleaching, and sticking. A correlation between the observed dynamic indicator and the combined residual anisotropy allowed us to experimentally validate the threshold criterion of both dyes $r_{c,\infty} < 0.25$ when performing accurate FRET measurements. We also present indications that, when artifacts can be excluded, smFRET allows the sensitive detection of small-scale conformational fluctuations on the Ångström level. Hence, FRET can be used to investigate the dynamic behavior of biomolecular complexes on a wide range of time scales and is a powerful tool for the coming age of dynamic structural biology.

Methods:

Methods, including statements of data availability and any associated accession codes and references, are available in the online methods.

Acknowledgements:

Work in the lab of T.C. was financed by an ERC Starting Grant (ERC-StG 638536—SM-IMPORT), Deutsche Forschungsgemeinschaft within GRK2062 (project C03) and SFB863 (project A13) and an Alexander von Humboldt postdoctoral fellowship (to N.Z.). T.C. and D.C.L acknowledge the support of the LMUexcellent and the Center for integrated protein science Munich (CiPSM). D.C.L. acknowledges the support of the Nanosystems Initiative Munich (NIM) and the LMU via the LMUinnovative program BioImaging Network (BIN). We also acknowledge support via the SFB1035 (German Research Foundation DFG, Sonderforschungsbereich 1035 Projektnummer 201302640, Project A11 to D.C.L. and Project B03 to M.Sa.). C.A.M.S. acknowledges the support by the European Research Council (ERC; grant No. 671208 (hybridFRET)) and by the Deutsche Forschungsgemeinschaft (DFG, grant SE 1195/17-1). Research in the contributing authors labs was financed by the following sources: P.T acknowledges the support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 201269156 – SFB 1032 (A13) and Project-ID 267681426. T.D.C was supported by the BBSRC (BB/T008032/1) and EPSRC (EP/V034804/1), B.A. was supported by an EPSRC Prize Fellowship. BMBF grant 03Z2EN11 and 03Z22E511 as well as DFG grant SCHL 1896/4-1 (to M.Sc.). SFB960 Project A7 (to D.G.) US National Institutes of Health grant GM122569 (to T.Ha). J.H. acknowledges the Research Foundation Flanders (FWO, projects G0B4915, G0B9922N and G0H3716N) and is indebted to Johan Hofkens at KU Leuven for the used smFRET infrastructure. ERC grant

agreement no. 681891 (Prosint) and German Research Foundation (DFG) under Germany's Excellence Strategy (CIBSS EXC-2189 Project ID 390939984) (to T.Hu.). Royal Society Dorothy Hodgkin Research Fellowship DKR00620 and Research Grant for Research Fellows RGF\R1\180054 (to N.C.R.), by the Wellcome Trust (110164/Z/15/Z to A.N.K.). The Israel Science Foundation (grant 3565/20 to E.L., within the KillCorona – Curbing Coronavirus Research Program), the NIH (grant R01 GM130942 to S.W. and to E.Ler. as a subaward), by the Milner Fund (to E.Ler.), and by the Hebrew University of Jerusalem (start-up funds to E.Ler.). Agence Nationale pour la Recherche (ANR 18-CE11-0004-02, ANR-19-CE44-0009-02, ANR-21-CE11-0034-01, ANR-21-CE11-0026-03 and ANR-10-INBS-04, “Investments for the future” to E.M.)].

Competing interests:

The authors declare no competing interests.

Author contributions:

C.G. and T.C. initiated the study. C.G., G.A., A.B., C.A.M.S., D.C.L. and T.C. designed research. A.B., C.A.M.S., D.C.L. and T.C. supervised the project. R.M. cloned and purified MalE. H.S.K. and M.S. provided U2AF2. C.G. and G.A. performed labeling of MalE and U2AF2 variants for shipment to participating labs. M.P., J.F. and C.A.M.S. designed MalE mutants 4 and 5 in silico. M.P. and J.F. performed initial measurements on MalE mutants 4 and 5. G.A. reperformed the analysis on the provided raw data for U2AF2 and MalE-1 from 8 laboratories. G.A., A.B. and M.P. performed dynamic shift estimation. C.G. performed FCS experiments on MalE variants, time-resolved anisotropy experiments and R_0-determination. M.P. performed time-resolved anisotropy analysis of single labelled MalE cysteine mutants from ensemble measurements as well of all MalE and U2AF2 dye combinations from smFRET measurements. C.G., G.A. and M.P. performed measurements of MalE mutants with additional dye combinations and M.P. and A.B. performed statistical analysis of dynamic shifts and anisotropies. G.A. and A.B. performed estimation of setup-dependent parameters and PDA of U2AF2. G.A. performed the filtered-FCS and A.B. performed TCSPC analysis of U2AF2. C.G., G.A. and M.P. performed smFRET measurements on dsDNA rulers. M.P. performed AV and ACV modelling of dye distributions for MalE and U2AF2. G.M. performed MST experiments. M.d.B. performed confocal scanning experiments for surface-immobilized MalE. All authors were involved in performing comparison experiments and analyzing smFRET data. C.G. and G.A. consolidated data collection of participating labs. C.G. and G.A. designed Fig. 1. C.G., G.A. and M.P. designed Fig. 2. A.B. designed Fig.s 3 and 4. A.B. and M.P. designed Figs. 5 and 6. C.G., G.A., M.P., A.B., C.A.M.S., D.C.L. and T.C. interpreted data and wrote the manuscript in consultation with all authors.

Data Availability:

The data for all figures, most supplementary figures, the raw data for MalE measurements from eight laboratories and the raw data for all U2AF2 measurements have been uploaded to Zenodo (https://doi.org/10.5281/zenodo.6977332). The remaining data will be uploaded before publication.

References:

1. Lerner, E. et al. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. eLife 10, e60416 (2021).

2. Lerner, E. et al. Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. Science 359, eaan1133 (2018).

3. Algar, W. R., Hildebrandt, N., Vogel, S. S. & Medintz, I. L. FRET as a biomolecular research tool — understanding its potential while avoiding pitfalls. Nat. Methods 16, 815–829 (2019).

4. Hildebrandt, N. How to Apply FRET: From Experimental Design to Data Analysis. FRET - Förster Reson. Energy Transf. From Theory to Appl. 105–163 (2013).

5. Muschielok, A. et al. A nano-positioning system for macromolecular structural analysis. Nat. Methods 5, 965–971 (2008).

6. Kalinin, S. et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat. Methods 9, 1218–1225 (2012).

7. Craggs, T. D. & Kapanidis, A. N. Six steps closer to FRET-driven structural biology. Nat. Methods 9, 1157–1159 (2012).

8. Voith von Voithenberg, L. & Lamb, D. C. Single Pair Förster Resonance Energy Transfer: A Versatile Tool To Investigate Protein Conformational Dynamics. BioEssays 40, 1–14 (2018).

9. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: Single-molecule FRET and beyond. Chem. Soc. Rev. 43, 1156–1171 (2014).

10. Krainer, G., Hartmann, A. & Schlierf, M. FarFRET: Extending the Range in Single-Molecule FRET Experiments beyond 10 nm. Nano Lett. 15, 5826–5829 (2015).

11. Muschielok, A. & Michaelis, J. Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks. J. Phys. Chem. B 115, 11927–11937 (2011).

12. Sali, A. et al. Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. Structure 23, 1156–1167 (2015).

13. Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M. & Hugel, T. Multidomain structure and correlated dynamics determined by self-consistent FRET networks. Nat. Methods 14, 176–182 (2017).

14. Choi, U. B. et al. Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex. Nat. Struct. Mol. Biol. 17, 318–324 (2010).

15. Dimura, M. et al. Automated and optimally FRET-assisted structural modeling. Nat. Commun. 11, 1–14 (2020).

16. Lerner, E., Ingargiola, A. & Weiss, S. Characterizing highly dynamic conformational states: The transcription bubble in RNAP-promoter open complex as an example. J. Chem. Phys. 148, 123315 (2018).

17. Craggs, T. D. et al. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. Nucleic Acids Res. 47, 10788–10800 (2019).

18. Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. Nat. Methods 15, 669–676 (2018).

19. Rout, M. P. & Sali, A. Principles for Integrative Structural Biology Studies. Cell 177, 1384–1403 (2019).

20. Sali, A. From integrative structural biology to cell biology. J. Biol. Chem. 296, 100743, 1–15 (2021).

21. Burley, S. K. et al. PDB-Dev: a Prototype System for Depositing Integrative/Hybrid Structural Models. Structure 25, 1317–1318 (2017).

22. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. Microbiol. Mol. Biol. Rev. 72, 317–364 (2008).

23. Mächtel, R., Narducci, A., Griffith, D. A., Cordes, T. & Orelle, C. An integrated transport mechanism of the maltose ABC importer. Res. Microbiol. 170, 321–337 (2019).

24. Malik, A. Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of E. coli. 3 Biotech 6, 1–7 (2016).

25. Berntsson, R. P. A., Smits, S. H. J., Schmitt, L., Slotboom, D. J. & Poolman, B. A structural classification of substrate-binding proteins. FEBS Lett. 584, 2606–2617 (2010).

26. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. Domain dislocation: A change of core structure in periplasmic binding proteins in their evolutionary history. J. Mol. Biol. 286, 279–290 (1999).

27. Banerjee, H., Rahn, A., Davis, W. & Singh, R. Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF65) recognize polypyrimidine tracts using multiple modes of binding. Rna 9, 88–99 (2003).

28. Sickmier, E. A. et al. Structural Basis for Polypyrimidine Tract Recognition by the Essential Pre-mRNA Splicing Factor U2AF65. Mol. Cell 23, 49–59 (2006).

29. Huang, J. R. et al. Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: A combined NMR and SAXS study. J. Am. Chem. Soc. 136, 7068–7076 (2014).

30. MacKereth, C. D. et al. Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. Nature 475, 408–413 (2011).

31. Kapanidis, A. N. et al. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. Proc. Natl. Acad. Sci. U. S. A. 101, 8936–8941 (2004).

32. Kapanidis, A. N. et al. Alternating-laser excitation of single molecules. Acc. Chem. Res. 38, 523–533 (2005).

33. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. Biophys. J. 89, 3508–3522 (2005).

34. Lee, N. K. et al. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. Biophys. J. 88, 2939–2953 (2005).

35. Kudryavtsev, V. et al. Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. ChemPhysChem 13, 1060–1078 (2012).

36. Von Voithenberg, L. V. et al. Recognition of the 3′ splice site RNA by the U2AF heterodimer involves a dynamic population shift. Proc. Natl. Acad. Sci. U. S. A. 113, E7169–E7175 (2016).

37. Sánchez-Rico, C., Voith von Voithenberg, L., Warner, L., Lamb, D. C. & Sattler, M. Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by spFRET and NMR Spectroscopy. Chem. - A Eur. J. 23, 14267–14277 (2017).

38. Eggeling, C., Widengren, J., Rigler, R. & Seidel, C. A. M. Photobleaching of Fluorescent Dyes under Conditions Used for Single-Molecule Detection: Evidence of Two-Step Photolysis. Anal. Chem. 70, 2651–2659 (1998).

39. Chung, H. S., McHale, K., Louis, J. M. & Eaton, W. A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. Science 335, 981–984 (2012).

40. Ramanathan, R. & Muñoz, V. A Method for extracting the free energy surface and conformational dynamics of fast-folding proteins from single molecule photon trajectories. J. Phys. Chem. B 119, 7944–7956 (2015).

41. McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophys. J. 91, 1941–1951 (2006).

42. Liu, Y., Park, J., Dahmen, K. A., Chemla, Y. R. & Ha, T. A comparative study of multivariate and univariate hidden Markov modelings in time-binned single-molecule FRET data analysis. J. Phys. Chem. B 114, 5386–5403 (2010).

43. Bronson, J. E., Fei, J., Hofman, J. M., Gonzalez, R. L. & Wiggins, C. H. Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. Biophys. J. 97, 3196–3205 (2009).

44. Margittai, M. et al. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. Proc. Natl. Acad. Sci. U. S. A. 100, 15516–15521 (2003).

45. Diez, M. et al. Proton-powered subunit rotation in single membrane-bound F 0F1-ATP synthase. Nat. Struct. Mol. Biol. 11, 135–141 (2004).

46. Torres, T. & Levitus, M. Measuring conformational dynamics: A new FCS-FRET approach. J. Phys. Chem. B 111, 7392–7400 (2007).

47. Felekyan, S., Sanabria, H., Kalinin, S., Kühnemuth, R. & Seidel, C. A. M. Analyzing Förster resonance energy transfer with fluctuation algorithms. in Methods in Enzymology (ed. Tetin, S. Y. B. T.-M. in E.) 519, 39–85 (Academic Press, 2013).

48. Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A. & Seidel, C. A. M. Filtered FCS: Species auto- and cross-correlation functions highlight binding and dynamics in biomolecules. ChemPhysChem 13, 1036–1053 (2012).

49. Olofsson, L. et al. Fine tuning of sub-millisecond conformational dynamics controls metabotropic glutamate receptors agonist efficacy. Nat. Commun. 5, 5206 (2014).

50. Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying molecular dynamics in single-molecule fret experiments with burst variance analysis. Biophys. J. 100, 1568–1577 (2011).

51. Tomov, T. E. et al. Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. Biophys. J. 102, 1163–1173 (2012).

52. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. M. Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. J. Phys. Chem. B 114, 7983–7995 (2010).

53. Gopich, I. V. & Szabo, A. Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET. Proc. Natl. Acad. Sci. U. S. A. 109, 7747–7752 (2012).

54. Nettels, D., Gopich, I. V., Hoffmann, A. & Schuler, B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. Proc. Natl. Acad. Sci. U. S. A. 104, 2655–2660 (2007).

55. Hoffmann, A. et al. Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: Recurrence analysis of single particles (RASP). Phys. Chem. Chem. Phys. 13, 1857–1871 (2011).

56. Gopich, I. V. & Szabo, A. Decoding the pattern of photon colors in single-molecule FRET. J. Phys. Chem. B 113, 10965–10973 (2009).

57. Chung, H. S. & Gopich, I. V. Fast single-molecule FRET spectroscopy: Theory and experiment. Phys. Chem. Chem. Phys. 16, 18644–18657 (2014).

58. Pirchi, M. et al. Photon-by-Photon Hidden Markov Model Analysis for Microsecond Single-Molecule FRET Kinetics. J. Phys. Chem. B 120, 13065–13075 (2016).

59. Harris, P. D. et al. Multi-parameter photon-by-photon hidden Markov modeling. Nat. Commun. 13, 1000 (2022).

60. Ingargiola, A., Weiss, S. & Lerner, E. Monte Carlo Diffusion-Enhanced Photon Inference: Distance Distributions and Conformational Dynamics in Single-Molecule FRET. J. Phys. Chem. B 122, 11598–11615 (2018).

61. Barth, A. et al. Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. J. Chem. Phys. 156, 141501 (2022).

62. De Boer, M. et al. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. eLife 8, e44652 (2019).

63. Dimura, M. et al. Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. Current Opinion in Structural Biology 40, 163–185 (2016).

64. Beckers, M., Drechsler, F., Eilert, T., Nagy, J. & Michaelis, J. Quantitative structural information from single-molecule FRET. Faraday Discuss. 184, 117–129 (2015).

65. Sindbert, S. et al. Accurate distance determination of nucleic acids via Förster resonance energy transfer: Implications of dye Linker length and rigidity. J. Am. Chem. Soc. 133, 2463–2480 (2011).

66. Steffen, F. D., Sigel, R. K. O. & Börner, R. An atomistic view on carbocyanine photophysics in the realm of RNA. Phys. Chem. Chem. Phys. 18, 29045–29055 (2016).

67. Klose, D. et al. Resolving distance variations by single-molecule FRET and EPR spectroscopy using rotamer libraries. Biophys. J. 120, 4842–4858 (2021).

68. Reinartz, I. et al. Simulation of FRET dyes allows quantitative comparison against experimental data. J. Chem. Phys. 148, 1–9 (2018).

69. Hoefling, M. et al. Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and monte carlo approach. PLoS One 6, 19791 (2011).

70. Peulen, T. O., Opanasyuk, O. & Seidel, C. A. M. Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. J. Phys. Chem. B 121, 8211–8241 (2017).

71. Dale, R. E., Eisinger, J. & Blumberg, W. E. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. Biophys. J. 26, 161–193 (1979).

72. Dale, R. E. & Eisinger, J. Intramolecular distances determined by energy transfer. Dependence on orientational freedom of donor and acceptor. Biopolymers 13, 1573–1605 (1974).

73. Ivanov, V., Li, M. & Mizuuchi, K. Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. Biophys. J. 97, 922–929 (2009).

74. Eilert, T., Kallis, E., Nagy, J., Röcker, C. & Michaelis, J. Complete Kinetic Theory of FRET. J. Phys. Chem. B 122, 11677–11694 (2018).

75. Maier, J. A. et al. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput. 11, 3696–3713 (2015).

76. Zaccai, G. How soft is a protein? A protein dynamics force constant measured by neutron scattering. Science 288, 1604–1607 (2000).

77. Vinogradov, S. A. & Wilson, D. F. Recursive maximum entropy algorithm and its application to the luminescence lifetime distribution recovery. Appl. Spectrosc. 54, 849–855 (2000).

78. Ingargiola, A., Lerner, E., Chung, S. Y., Weiss, S. & Michalet, X. FRETBursts: An open source toolkit for analysis of freely-diffusing Single-molecule FRET. PLoS One 11, 39198 (2016).

79. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. Biophys. J. 114, 1518–1528 (2018).

80. Ambrose, B. et al. The smfBox is an open-source platform for single-molecule FRET. Nat. Commun. 11, 5641 (2020).

81. Knight, J. L., Mekler, V., Mukhopadhyay, J., Ebright, R. H. & Levy, R. M. Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: Developing benchmarks of model quality and reliability. Biophys. J. 88, 925–938 (2005).

82. Kapanidis, A. N. et al. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. Science 314, 1144–1147 (2006).

83. Sanabria, H. et al. Resolving dynamics and function of transient states in single enzyme molecules. Nat. Commun. 11, 1231 (2020).

84. Gouridis, G. et al. Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. Nat. Struct. Mol. Biol. 22, 57–64 (2015).

85. Jerabek-Willemsen, M. et al. MicroScale Thermophoresis: Interaction analysis and beyond. J. Mol. Struct. 1077, 101–113 (2014).

86. Hall, J. A., Gehring, K. & Nikaido, H. Two modes of ligand binding in maltose-binding protein of Escherichia coli: Correlation with the structure of ligands and the structure of binding protein. J. Biol. Chem. 272, 17605–17609 (1997).

87. Kim, E. et al. A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. Nat. Chem. Biol. 9, 313–318 (2013).

88. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection. Methods Enzymol. 475, 455–514 (2010).

89. Eggeling, C. et al. Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. J. Biotechnol. 86, 163–180 (2001).

Online Methods

Sample preparation of proteins. Double-cysteine mutants of MalE were prepared and labeled using established protocols[62]. Human RRM1,2 L187C-G326C mutant (U2AF2-148-342) was obtained and purified as described in Mackereth et al.[30].

Fluorescence labeling of proteins. All fluorophores were purchased as maleimide derivatives from commercial suppliers as listed in Supplementary Table 19. MalE was stochastically labeled as described previously[84] with fluorophores as indicated in the text with a combined labeling efficiency higher than 70% resulting in a donor-acceptor pairing of at least 20%. Protein stability and functionality (ligand binding) was verified by affinity measurements using microscale thermophoresis[85]. All preparations, i.e., MalE-wildtype, unlabeled cysteine mutants and fluorophore-labeled variants, showed an affinity for maltose between ~1-2 µM (Supplementary Fig. 5) consistent with previously published K_d-values for wild type MalE[86,87]. The stability and labeling of the sample were verified by fluorescence correlation spectroscopy (Supplementary Fig. 18), which excluded the presence of larger aggregates in the samples and confirms that MalE is functional.

U2AF2 was stochastically labeled as described previously in Voith von Voithenberg et al.[36]. The combined labeling efficiency for labeling reactions were 20% and 14% for Alexa546-Alexa647 and Atto532-Atto643 pairs, respectively. For Alexa488-Alexa647, the combined labeling efficiency was found to be 10%. The functionality of the labeled U2AF protein was checked with affinity measurement for U9 RNA, which was found to be 1.2 µM[30], consistent with the previous reports[36] (Supplementary Fig. 7d).

Sample handling. Both protein systems required special handling due to sample instability or aggregate formation, which are both problematic for long-term storage and shipping. The labeled MalE proteins were stored in 50 mM Tris-HCl pH 7.4, 50 mM KCl with 1 mg/ml bovine serum albumin (BSA) at 4°C for less than 7 days. U2AF2 was stored in 20 mM potassium phosphate buffer pH 6.5, 50 mM NaCl and kept in the fridge until used. Both samples were loaded in low-binding Eppendorf tubes (Eppendorf Germany, Catalog number 0030108094) and shipped on ice in a cooling box with overnight shipping to avoid unnecessary freezing and thawing. MalE stock solutions were on the order of 10 to 100 nM concentration and the sent stock solution of U2AF2 was 5-10 µM concentration. Dilution buffer for apo and holo measurement were provided. SmFRET experiments were carried out by diluting the labeled proteins to concentrations of ≈50 pM in 50 mM Tris-HCl pH 7.4, 50 mM KCl supplemented with the ligand maltose at 1 mM concentration. Labeled U2AF2 protein was measured at ~40-100 pM in 20 mM potassium phosphate buffer pH 6.5, 50 mM NaCl. Purchased U9 RNA (Biomers.net GmbH, Ulm, Germany, IBA Solutions for Life Sciences, Göttingen, Germany) was dissolved in RNA-free water and added directly to the solution at a final concentration of 5 µM for the holo measurements. Both proteins were studied on coverslips typically passivated with 1 mg/ml BSA in buffer before adding the

sample. The measurements were performed without any photo-stabilizer to keep the measurements as simple as possible to avoid any further source for discrepancies between the groups, e.g., degradation of photostabilizer or use of different photostabilizer concentrations.

SmFRET data acquisition and analysis. Data acquisition and correction procedures were performed for confocal measurements as described by Hellenkamp *et al.*[18]. The samples were measured using alternating laser excitation mode (ALEX) or Pulsed Interleaved Excitation (PIE) on a confocal microscope as sketched in Supplementary Fig. 2. A description of experimental procedures of all labs are given in Supplementary Note 18.

Briefly, the three recorded intensity time traces for each single-molecule event are:

donor emission after donor excitation: ${}^{\mathrm{i}}I_{\mathrm{Dem|Dex}}$,
acceptor emission after donor excitation (FRET signal): ${}^{\mathrm{i}}I_{\mathrm{Aem|Dex}}$,
and acceptor emission after acceptor excitation: ${}^{\mathrm{i}}I_{\mathrm{Aem|Aex}}$.

The apparent (raw) FRET efficiency is computed as:

$$E_{\mathrm{app}} = \frac{{}^{\mathrm{i}}I_{\mathrm{Aem|Dex}}}{{}^{\mathrm{i}}I_{\mathrm{Dem|Dex}} + {}^{\mathrm{i}}I_{\mathrm{Aem|Dex}}}, \tag{3}$$

Recorded intensities were corrected for background contributions as:

$${}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} = {}^{\mathrm{i}}I_{\mathrm{Dem|Dex}} - I^{(\mathrm{BG})}_{\mathrm{Dem|Dex}}, \tag{4}$$

$${}^{\mathrm{ii}}I_{\mathrm{Aem|Dex}} = {}^{\mathrm{i}}I_{\mathrm{Aem|Dex}} - I^{(\mathrm{BG})}_{\mathrm{Aem|Dex}}, \tag{5}$$

$${}^{\mathrm{ii}}I_{\mathrm{Aem|Aex}} = {}^{\mathrm{i}}I_{\mathrm{Aem|Aex}} - I^{(\mathrm{BG})}_{\mathrm{Aem|Aex}}, \tag{6}$$

where $I^{(\mathrm{BG})}_{\mathrm{Dem|Dex}}$, $I^{(\mathrm{BG})}_{\mathrm{Aem|Dex}}$, and $I^{(\mathrm{BG})}_{\mathrm{Aem|Aex}}$ are the respective background signals. Correction factors for spectral crosstalk, α, and direct excitation, δ, were determined from the donor-only and acceptor-only populations[34]. The corrected acceptor fluorescence after donor excitation, $F_{A|D}$, is computed as:

$$F_{A|D} = {}^{\mathrm{ii}}I_{\mathrm{Aem|Dex}} - \alpha\,{}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} - \delta\,{}^{\mathrm{ii}}I_{\mathrm{Aem|Aex}} \tag{7}$$

The γ and β factors, correcting for differences in the detection yield and excitation fluxes of the donor and acceptor dyes, were estimated using a global correction procedure using the approach of Lee *et al*. (Supplementary Fig. 3)[34]. Alternatively, when pulsed excitation was used and the sample is known to be static, the γ factor can be determined by fitting the measured population to the static FRET line[35,88]. This allows a good determination of the γ factor when only a single species is present but requires a static sample and the appropriate static FRET line (Supplementary Note 2).

The accurate FRET efficiency *E* and stoichiometry *S* values were then calculated as:

$$E = \frac{F_{A|D}}{\gamma^{\mathrm{ii}} I_{\mathrm{Dem|Dex}} + F_{A|D}}, \tag{8}$$

$$S = \frac{\gamma^{\mathrm{ii}} I_{\mathrm{Dem|Dex}} + F_{A|D}}{\gamma^{\mathrm{ii}} I_{\mathrm{Dem|Dex}} + F_{A|D} + {}^{\mathrm{ii}} I_{\mathrm{Aem|Aex}}/\beta}. \tag{9}$$

Conversion of accurate FRET efficiencies into distances were done using Eq. 2 with Förster radii determined as described in Supplementary Note 7.

Detection of Protein Dynamics. In this work, we used the following two approaches to detect conformational dynamics:

Burst Variance Analysis (BVA):

In BVA, the presence of dynamics is determined by looking for excess variance in the FRET efficiency data beyond the shot-noise limit. The the standard deviation ($\sigma_{E_{\mathrm{app}}}$) of the apparent FRET efficiency (E_{app}) is calculated using a fixed photon window of 5 (n) over the time period of the individual bursts given by:

$$\sigma_{E_{\mathrm{app}}} = \sqrt{\frac{E_{\mathrm{app}}\left(1 - E_{\mathrm{app}}\right)}{n}}, \tag{10}$$

The shot-noise limited standard deviation of the apparent FRET efficiency is generally described by a semi-circle[50] (Fig. 4a and Supplementary Fig. 11a-d). In the presence of dynamics, the standard deviation for the FRET efficiency within a burst becomes higher than that expected from shot noise. Photophysical effects like bleaching and blinking also give rise to the higher standard deviation beyond the shot-noise limit. Typically, BVA is sensitive to fluctuations in FRET signal of $\gtrsim$ 100 µs, but these depends on the brightness of the burst and the photon window used.

FRET efficiency versus fluorescence-weighted average donor lifetime analysis (E-τ plots):

Two-dimensional histograms of the FRET efficiency E and donor fluorescence lifetime $\langle \tau_{D(A)} \rangle_F$ (Fig. 4b and Supplementary Fig. 11e-h) were created for single molecule measurements using MFD in combination with pulsed-interleaved excitation (PIE)[35], described below. Static FRET lines were calculated using the following equation:

$$E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}} \tag{11}$$

and further modified for linker dynamics[61]. Deviations of FRET populations from the static FRET line can indicate FRET dynamics, which can be due to conformational fluctuations or photophysical dynamics. In addition, a time-resolved FRET analysis of TCSPC data can accurately resolve the distance heterogeneities by revealing multiple components in the decay of the curve and recovers their specific species fractions and FRET rate constants[70]. Dynamics are thus detected from the presence of multiple components in the sub-ensemble decay of a single FRET population. In addition, dynamics that are slower than the fluorescence lifetime (~ 5 ns) are

not averaged in the FRET lifetime analysis leading to detection of the full conformational distribution.

Multiparameter fluorescence detection (MFD) with Pulsed Interleaved Excitation (PIE).

MFD, introduced by Eggeling et al.[89], combines spectral and polarized detection with picosecond pulsed lasers and time correlated single photon counting (TCSPC), allowing the simultaneous detection of intensity, lifetime, anisotropy and spectral range of the fluorescence signal of single molecules. nsALEX of PIE additionally provide the acceptor lifetime information[35]. Due to the availability of the lifetime information when using pulsed excitation, this approach is well suited for using *E*-τ-based analyses.

AV simulations: The AV approach employs a simple coarse-grained dye model[65] defined by five parameters: the width and length of the linker, and three radii that define the fluorophore volume (Fig. 5a, Supplementary Table 10). Using these parameters, AV simulations for both fluorophores were calculated by considering the linker flexibility and steric hindrances of the labeled molecule (Fig. 5a). In the ACV model[63], the residual anisotropy was used to estimate the fraction of sticking dyes. In the computation of the FRET-averaged model distances, the occupancy of a thin surface layer (~3 Å) was then increased such that its fraction matches the amount of interacting dye detected in the experiment (Fig. 5b Supplementary Table 10).

Supplementary Information:

Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins

Gebhardt, Agam, Popara, et al.

Supplementary Notes

Supplementary Note 1: Comparison between µsALEX and PIE

With the lifetime information, it is possible to directly access distances and distance distributions via analysis of the donor fluorescence decay[1,2] and detect dynamics on the nano- to millisecond timescale in the E-τ plot, as we have shown for the apo state of U2AF65. The advantage of µsALEX is that it is less expensive to implement than nsALEX/PIE and that the continuous wave excitation exerts less photophysical stress on the fluorophores (e.g., bleaching or blinking) compared to the high peak irradiance for pulsed excitation[3]. On the other hand, nsALEX/PIE has the advantage that they provide the fluorescence lifetime information of the donor and acceptor fluorophore[4]. Via the lifetime information, it is also possible to detect changes of the donor and acceptor quantum yields due to protein or ligand interactions with the dye. In addition, the faster alternation timescale of nsALEX/PIE enables higher time-resolution for fluorescence correlation spectroscopy and reduces the spread in the stoichiometry distribution due to diffusion of molecules through the confocal volume during the alternation period. Combining pulsed excitation with polarized excitation and detection[4,5] further adds the fluorescence anisotropy information on the single-molecule or sub-ensemble level, allowing one to monitor size changes of the biomolecule (e.g., through the binding of interaction partners) and detect changes of the local environment of the fluorophore such as dye sticking (as seen for MalE-1) directly from the single-molecule dataset.

The spread in the results in Fig. 3d indicate that the direct probing of the acceptor in ALEX/PIE experiments was applied differently by the participating labs. Part of the participants either used high laser power for acceptor excitation to obtain reliable information on the labeling stoichiometry, acceptor lifetime and anisotropy by acquiring a maximum number of photons. Others kept the acceptor excitation power as low as possible to minimize acceptor saturation and photobleaching[6]. Regardless of which approach was used, accurate FRET efficiencies can be determined from the experimentally determined correction factors. Acceptor photobleaching results in a higher amount of donor only molecules. In addition, acceptor saturation leads to dark states that still quench the donor via FRET and appears below the static FRET line. Acceptor blinking leads to a mixture of donor-only signal, which results in false-positive dynamics.

Supplementary Note 2: Primer design and sequences for creation of MalE variants

The sequence of MalE (blue, DNA sequence in black) and the primers (in red) used for generating the mutants are given below. The procedures have been published previously[7].

```
1                                                    GTCGGTAAG
1                 K  I  E  E  G  K  L  V  I  W  I  N  G  D  K  G  Y  N  G  L  A  E  V  G  K
1 AAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAG
```

```
  76 AAATTCGAGWRMGATACCGG 3' Lys29Cys
  26                   K  F  E  K  D  T  G  I  K  V  T  V  E  H  P  D  K  L  E  E  K  F  P  Q  V
  76 AAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTT

  51                   A  A  T  G  D  G  P  D  I  I  F  W  A  H  D  R  F  G  G  Y  A  Q  S  G  L
 151 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTG

 226                CCGGACAAAGCGTTCCAGKRCAAGCTGTATCCG 3' Asp87Cys
  76                   L  A  E  I  T  P  D  K  A  F  Q  D  K  L  Y  P  F  T  W  D  A  V  R  Y  N
 226 TTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAAC

 101                   G  K  L  I  A  Y  P  I  A  V  E  A  L  S  L  I  Y  N  K  D  L  L  P  N  P
 301 GGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCG

 376             GAAGAGATCCCGKSSCTGGATAAAGAAC 3' Ala134Cys
 126                   P  K  T  W  E  E  I  P  A  L  D  K  E  L  K  A  K  G  K  S  A  L  M  F  N
 376 CCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC

 151                   L  Q  E  P  Y  F  T  W  P  L  I  A  A  D  G  G  Y  A  F  K  Y  E  N  G  K
 451 CTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAG

 526                      GTGGATAACKSYGGCGCGAAAGCG 3' Ala186Cys
 176                   Y  D  I  K  D  V  G  V  D  N  A  G  A  K  A  G  L  T  F  L  V  D  L  I  K
 526 TACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAA

 201                   N  K  H  M  N  A  D  T  D  Y  S  I  A  E  A  A  F  N  K  G  E  T  A  M  T
 601 AACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACC

 226                   I  N  G  P  W  A  W  S  N  I  D  T  S  K  V  N  Y  G  V  T  V  L  P  T  F
 676 ATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTC

 251                   K  G  Q  P  S  K  P  F  V  G  V  L  S  A  G  I  N  A  A  S  P  N  K  E  L
 751 AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTG

 276                   A  K  E  F  L  E  N  Y  L  L  T  D  E  G  L  E  A  V  N  K  D  K  P  L  G
 826 GCGAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGT

 301                   A  V  A  L  K  S  Y  E  E  E  L  A  K  D  P  R  I  A  A  T  M  E  N  A  Q
 901 GCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCCAG

 976                                                                  GATCAACGCC
 326                   K  G  E  I  M  P  N  I  P  Q  M  S  A  F  W  Y  A  V  R  T  A  V  I  N  A
 976 AAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC

1051 GCCWGCGGTCGTCAG 3' Ser352Cys
 351                                        A  S  G  R  Q  T  V  D  E  A  L  K  D  A  Q  T
1051 GCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT
```

Supplementary Note 3: Discussion of γ-factor estimation.

Obtaining a reliable γ-factor is very crucial for smFRET data analysis and proper determination of accurate FRET values. While correction for background, spectral crosstalk and direct excitation can be performed reliably, the detection efficiency correction factor γ is much more difficult to determine and the best approach depends on the sample at hand and the measurement modality (see Supplementary Note 1, Supplementary Note 19). A wrong γ-factor introduces systematic errors in the FRET efficiency and the derived distances, especially for intermediate FRET efficiencies. The difficulties arising from the γ factor can be illustrated with data from one lab that used sub-optimal filter combinations for the dye pair Alexa546/Alexa647. This resulted in an inefficient detection of the red fluorescent signal and a γ-factor of 0.09 (Supplementary Table 1; lab #18). Consequently, the determined accurate FRET efficiencies were unreliable, i.e., deviating largely from the expected values, and were hence excluded for the calculation of the averages in Fig. 1. For another laboratory, the correction factors could not be determined due to a missing or non-functional red laser (Supplementary Table 1; lab #19). Hence, accurate FRET values could not be calculated. A wrong γ-factor can additionally distort the dynamic shift in the E-τ plot resulting either in unphysically negative or artificially positive apparent dynamic shifts. The absence of negative values for the apparent dynamic shifts in Fig. 4f indicates that the five participating groups estimated the γ-factor well. Note that no specific instructions were given on how the selection of the donor-only, acceptor-only and donor-acceptor labeled subpopulations for the determination of α, β, γ and δ parameters should be performed.

In this study, we asked all groups to use a global γ-factor for the data analysis of MalE, meaning that one common γ-factor was determined for all six data sets (Supplementary Figure 3). A global γ-factor works well when all samples have the same photophysical behavior. However, when the labeling position and/or conformation of the biomolecule induces a change of the fluorescence quantum yield of the fluorophore by local quenching or enhancement, a global γ factor is no longer strictly correct. Here, we observed variations of the fluorescence lifetimes of the dyes on the order of 5-10% (Supplementary Table 9). This indicates no large variations of the fluorescence quantum yields between the different samples and justifies the global γ correction. On the other hand, a local γ correction can only be done when multiple conformations with different FRET efficiencies are present in a single measurement, or by using the lifetime information when the sample is known to be static. The γ-factors reported in this study for the MalE system with dye pair Alexa546-Alexa647 were generally low in the range from 0.2 to 0.6, with an average of 0.39±0.12 (Supplementary Table 1). Correspondingly, the estimated error of the γ-factor of ~23% (Fig. 3e) is large even though the absolute error is on the order of ±0.07.

Supplementary Note 4: Data re-evaluation procedure for U2AF65

Reanalysis of the collected raw data of U2AF65 from different labs was performed using the PAM software[8] (Supplementary Table 16). First, a burst search was performed using an all photon burst search with a threshold of 50-100 photons per sliding time window of 500 µs depending on the

dataset. For one set of measurements, a lower threshold of 20 photons per 500 µs time window was necessary. To remove blinking and bleaching events, an ALEX-2CDE filter with a lower limit of 5 and an upper limit of 25 was used depending on the data set. Values may differ depending on the excitation intensities and sample concentrations used for the measurements[9]. After burst selection, background subtraction and correction for crosstalk and direct excitation were performed as discussed in the data analysis section. Briefly, for background subtraction, the background signal was obtained from the buffer measurement and subtracted it from the burst signal. Crosstalk and direct excitation corrections were performed by calculating the signal in the acceptor channel for donor only and acceptor only species respectively, after donor excitation and subtracting them from the burst signal. To determine the detection correction factor, we used the approach of Lee *et al.*[10] (i.e. fitting a line to $1/S_{PR}$ vs E_{PR}) as the sample is dynamic and a lifetime approach is not possible. The apo configuration shows a single, dynamically averaged population so that we had to combine data from both the holo and apo measurements. We verified that there was no significant change in quantum yield of the donor and acceptor fluorophores in the absence and presence of ligand by measuring the fluorescence lifetime of the donor-only species and the acceptor lifetime with direct excitation. However, we did observe an additional subpopulation with a higher stoichiometry value (Supplementary Figure 15a). The acceptor is slightly quenched in this population, but not enough to explain the stoichiometry shift. As the origins of this population is unclear and simple explanations are insufficient to describe the observed properties, this population was not incorporated in the calculation of the γ-value. The average values of S_{PR} and E_{PR} for the three populations (apo, holo – low FRET, holo – high FRET) were determined from the peak values of a 2D-Gaussian fit in the ES-histograms for the respective populations. From these peak values, a straight line was fit to the three data points of $1/S_{PR}$ vs E_{PR}. Reanalysis of the data by a single person led to a further improvement of the consistency between laboratories (Fig. 2d-e). Part of the discrepancy came from the fact that the individual labs used a global γ approach but did not compensate for the presence of the second population. Further reasons for the discrepancies in the measured data from different laboratories arise from the fact that the dynamics and RNA binding are temperature dependent (and temperature was not specified) and that the RNA concentration used in the holo experiments was insufficient to saturate protein binding leading to a mixture of apo and holo proteins in the "holo" measurements.

Supplementary Note 5: Error propagation of the uncertainty of the γ-factor on the FRET efficiency *E*

To support the hypothesis that the spread of the reported FRET efficiency E values for the different MalE mutants is caused by inaccuracies of the experimental calibration, we derive how the uncertainty of the detection efficiency correction factor γ propagates into the uncertainty of the measured FRET efficiency. Using the nomenclature introduced by Hellenkamp et al.[11] , the apparent FRET efficiency before γ-factor correction ${}^{\mathrm{iii}}E_{\mathrm{app}}$ is given by:

$$ {}^{\mathrm{iii}}E_{\mathrm{app}} = \frac{F_{A|D}}{{}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}} + F_{A|D}}, \tag{5.1} $$

where ${}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}}$ is the background-corrected donor intensity after donor excitation, and $F_{A|D}$ is the cross-talk and direct excitation corrected acceptor fluorescence after donor excitation, given by:

$$ F_{A|D} = {}^{\mathrm{ii}}I_{\mathrm{A_{em}|D_{ex}}} - \alpha\,{}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}} - \delta\,{}^{\mathrm{ii}}I_{\mathrm{A_{em}|A_{ex}}}. \tag{5.2} $$

Here, ${}^{ii}I_{\mathrm{A_{em}|D_{ex}}}$ is the background-corrected acceptor intensity after donor excitation, ${}^{\mathrm{ii}}I_{\mathrm{A_{em}|A_{ex}}}$ is the background-corrected acceptor intensity after acceptor excitation, and α and δ are the correction factor for donor crosstalk and acceptor direct excitation respectively.

The fully corrected FRET efficiency E is then given by:

$$ E = \frac{F_{A|D}}{\gamma\,{}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}} + F_{A|D}} = \left(\gamma\left(\frac{1}{{}^{\mathrm{iii}}E_{\mathrm{app}}} - 1\right) + 1\right)^{-1} \tag{5.3} $$

Using standard error propagation, the uncertainty of the FRET efficiency E due to the uncertainty of the γ-factor is given by:

$$ \Delta E = \left|\frac{\partial E}{\partial \gamma}\right| \Delta\gamma, \tag{5.4} $$

where $\Delta\gamma$ is the uncertainty in γ, and $|\partial E/\partial\gamma|$ is the partial derivative of E with respect to γ. The partial derivative is given by:

$$ \left|\frac{\partial E}{\partial \gamma}\right| = \frac{E(1-E)}{\gamma} $$

Thus, the propagated uncertainty is given by:

$$ \Delta E = E(1-E)\frac{\Delta\gamma}{\gamma} \tag{5.5} $$

This equation was fit to the experimental data in Fig. 3e of the main text, yielding an estimated uncertainty of the γ-factor calibration of $\Delta\gamma/\gamma = 0.23$.

To estimate the absolute (rather than the relative) error of γ from the data, we calculated $\Delta\gamma$ directly from the reported FRET efficiency E and γ-factor of each lab by rearranging Eq. 5.5:

$$\Delta\gamma = \gamma \frac{\Delta E}{E(1-E)} \tag{5.6}$$

From the obtained values for $\Delta\gamma$ for the different labs, we obtain an average value of $\Delta\gamma = 0.071 \pm 0.051$ (mean ± standard deviation).

Supplementary Note 6: Estimation of the experimental dynamic shift

In BVA, the standard deviation of the apparent FRET efficiency, $\sigma_{E_{\mathrm{app}}}$, is plotted against the apparent FRET efficiency, E_{app} (Fig. 4a). For the E-τ plot, usually the accurate FRET efficiency E is plotted against the intensity-averaged donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_F$ (Fig. 4b). For the estimation of the dynamic shift, $\langle\tau_{D(A)}\rangle_F$ is normalized with respect to the lifetime of the donor in the absence of the acceptor, $\tau_{D(0)}$, to constrain both axes to the interval of [0,1].

To determine the peak position of the single-molecule population, we fit the distribution of single molecule events in the E-$\langle\tau_{D(A)}\rangle_F/\tau_{D(0)}$ or BVA histograms using a superposition of M two-dimensional Gaussian distributions. Each Gaussian distribution is described by an amplitude A_i, a vector of the central coordinates $\boldsymbol{\mu_i}$, and a covariance matrix $\boldsymbol{\Sigma_i}$:

$$f(\boldsymbol{x}) = \sum_{i=1}^{M} A_i \cdot (2\pi)^{-1} |\boldsymbol{\Sigma_i}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_i})\boldsymbol{\Sigma_i}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_i})^{'}\right), \tag{6.1}$$

where $\boldsymbol{\mu_i}$ is a row vector of length 2, and $\boldsymbol{\Sigma}$ is a *2*-by-*2* symmetric matrix whose diagonal elements and non-diagonal elements are the variances and covariances, respectively. Depending on whether the analysis is applied to the E-τ or BVA plot, the coordinate row vector $\boldsymbol{x}$ is given by $\boldsymbol{x} = \left(\langle\tau_{D(A)}\rangle_F/\tau_{D(0)}, E\right)$ or $\boldsymbol{x} = \left(E_{\mathrm{app}}, \sigma_{E_{\mathrm{app}}}\right)$, respectively. The model parameters are determined using a maximum likelihood estimation. The required number of populations needed to adequately describe the data is determined visually. When more than one bivariate normal distributions ($M > 1$) is needed to describe the data, the position of the component with the highest amplitude is taken. After the position of the major population is determined, we estimate the dynamic shift as described below.

For BVA, the dynamic shift is defined as the distance of the population center along the vertical (σ_{BVA}) axis from the static FRET-line, given by:

$$\sigma_{E_{\mathrm{app}}} = \sqrt{\frac{E_{\mathrm{app}}\left(1-E_{\mathrm{app}}\right)}{n}}, \tag{6.2}$$

where n is the number of photons per window used to estimate $\sigma_{E_{\mathrm{app}}}$ (here, $n = 5$). To provide consistency with the procedure proposed by Torella et al.[12], we compute the dynamic shift for the main population based on the average value of $\sigma_{E_{\mathrm{app}}}$ over a FRET efficiency interval given by its

estimated mean and width, $[\mu_{E_{\mathrm{app}}} - w_{E_{\mathrm{app}}}, \mu_{E_{\mathrm{app}}} + w_{E_{\mathrm{app}}}]$, where $\mu_{E_{\mathrm{app}}}$ and $w_{E_{\mathrm{app}}}$ are the center position and width of the main population in the $\left(E_{\mathrm{app}}, \sigma_{E_{\mathrm{app}}}\right)$ plot.

For the E-τ plot, the dynamic shift is defined as the minimum distance from the peak of the population to the static FRET line, given by:

$$E = 1 - \frac{\langle \tau_{D(A)} \rangle_F}{\tau_{D(0)}}. \tag{6.3}$$

Generally, a linker correction for the static FRET-line is performed[13]. As a result, the static FRET-line is slightly curved. The dynamic shift is thus not necessarily the orthogonal distance to the static FRET line and was determined numerically. Lastly, the dynamic shift is assigned a positive or negative sign depending on whether the population lies above (positive) or below (negative) the static FRET line.

Supplementary Note 7: Prediction of the expected dynamic shifts

Analytical expressions were derived to predict the maximum expected dynamic shift for a dynamic exchange between two conformational states with FRET efficiencies $E^{(1)}$ and $E^{(2)}$ in the E-τ and BVA plots.

I. The dynamic shift in the E-τ plot

For the E-τ plot, i.e. a plot of the intensity averaged FRET efficiency E against the normalized intensity averaged donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_F/\tau_{D(0)}$, the maximum dynamic shift is given[14] :

$$ds_{(E-\tau)} = \frac{1}{\sqrt{2}}\left(\sqrt{1-E^{(1)}} - \sqrt{1-E^{(2)}}\right)^2. \tag{7.1}$$

This equation is derived by considering the maximum deviation between the ideal static FRET-line, given by:

$$E^{(\mathrm{stat})}_{(E-\tau)} = 1 - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}, \tag{7.2}$$

where $\langle\tau_{D(A)}\rangle_F$ and $\tau_{D(0)}$ are the intensity-weighted average donor fluorescence lifetimes in the presence and absence of the acceptor, respectively. Binary dynamic exchange between two limiting conformational states with FRET efficiencies $E^{(1)}$ and $E^{(2)}$ and the corresponding donor fluorescence lifetimes $\tau^{(1)}_{D(A)}$ and $\tau^{(2)}_{D(A)}$ is described by the dynamic FRET-line for exchange between two states[13]:

$$E^{(\mathrm{dyn})}_{(E-\tau)} = 1 - \frac{\tau^{(1)}_{D(A)}\tau^{(2)}_{D(A)}}{\tau_{D(0)}\left(\tau^{(1)}_{D(A)} + \tau^{(2)}_{D(A)} - \langle\tau_{D(A)}\rangle_F\right)}. \tag{7.3}$$

Using the relations:

$$\tau^{(i)}_{D(A)} = \tau_{D(0)}\left(1-E^{(i)}\right);\ \ i = 1,2 \tag{7.4}$$

we can then write Eq. 7.3 as a function of the FRET efficiencies:

$$E^{(\mathrm{dyn})}_{(E-\tau)} = 1 - \frac{\left(1-E^{(1)}\right)\left(1-E^{(2)}\right)}{\left(2 - E^{(1)} - E^{(2)} - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}\right)} \tag{7.5}$$

The difference between the static and dynamic FRET-lines along the FRET efficiency axis, Δ_E, as a function of $\langle\tau_{D(A)}\rangle_F$ is then given by:

$$\Delta_E = E^{(\mathrm{dyn})}_{(E-\tau)} - E^{(\mathrm{stat})}_{(E-\tau)} \tag{7.6}$$

$$\Delta_E(\langle\tau_{D(A)}\rangle_F) = \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}} - \frac{\left(1-E^{(1)}\right)\left(1-E^{(2)}\right)}{\left(2-E^{(1)}-E^{(2)}-\frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}\right)} \tag{7.7}$$

The function $\Delta_E(\langle\tau_{D(A)}\rangle_F)$ is unimodal and the maximum depends on the FRET efficiencies of the limiting states, $E^{(1)}$ and $E^{(2)}$:

$$\Delta_{E,max} = \left(\sqrt{1-E^{(1)}} - \sqrt{1-E^{(2)}}\right)^2. \tag{7.8}$$

We define the *dynamic shift*, $ds_{(E-\tau)}$, as the maximum deviation of the dynamic FRET-line measured orthogonal to the static FRET-line (see Fig. 4b), which introduces the factor of $1/\sqrt{2}$:

$$ds_{(E-\tau)} \overset{\text{def}}{=} \frac{\Delta_{E,max}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\left(\sqrt{1-E^{(1)}} - \sqrt{1-E^{(2)}}\right)^2. \tag{7.9}$$

II. The dynamic shift in BVA

For the BVA plot, the maximum dynamic shift is defined as the maximum distance of the dynamic FRET-line from the static FRET-line along the y-axis, i.e., the estimated standard deviation of the apparent FRET efficiency, $\sigma_{E_{\mathrm{app}}}$ (see Fig. 4a). The static FRET-line in BVA, describing the shot noise variance, is given by[12] :

$$\sigma^{(\mathrm{stat})}_{E_{\mathrm{app}}} = \sqrt{\frac{E_{\mathrm{app}}(1-E_{\mathrm{app}})}{n}}, \tag{7.10}$$

where n is the photon averaging window used to estimate the standard deviation. For the dynamic FRET-line in BVA describing the exchange between two limiting conformational states with apparent FRET efficiencies $E^{(1)}_{\mathrm{app}}$ and $E^{(2)}_{\mathrm{app}}$, we need to consider the excess variance caused by the conformational exchange. The contributions of shot noise, described by Eqn. 7.10, and the conformational exchange are not additive. The variance of the signal does not depend on the time dependence of the fluctuations. Thus, it is not important to know at what time points transitions between the limiting states occurred. It is only important what fraction of time the molecule spent in each individual state. For a two-state system, the expected variance in the presence of shot-noise and conformational dynamics is given by:

$$\begin{aligned}\mathrm{Var}^{(\mathrm{dyn})}\left(E_{\mathrm{app}}\right) &= f_1\left[\frac{E_{\mathrm{app}}^{(1)}\left(1-E_{\mathrm{app}}^{(1)}\right)}{n}+{E_{\mathrm{app}}^{(1)}}^2\right]+f_2\left[\frac{E_{\mathrm{app}}^{(2)}\left(1-E_{\mathrm{app}}^{(2)}\right)}{n}+{E_{\mathrm{app}}^{(2)}}^2\right] \\ &\quad -\left(f_1E_{\mathrm{app}}^{(1)}+f_2E_{\mathrm{app}}^{(2)}\right)^2,\end{aligned} \tag{7.11}$$

where f_1 and f_2 are the fraction of time spent in the respective state with $f_1 + f_2 = 1$, and the apparent FRET efficiency is given by $E_{\mathrm{app}} = f_1E_{\mathrm{app}}^{(1)} + f_2E_{\mathrm{app}}^{(2)}$. The standard deviation of the apparent FRET efficiency for a dynamic exchange is then given by:

$$\sigma_{E_{\mathrm{app}}}^{(\mathrm{dyn})} = \sqrt{\mathrm{Var}^{(\mathrm{dyn})}\left(E_{\mathrm{app}}\right)}. \tag{7.12}$$

The dynamic FRET-line in BVA is then obtained by varying the fraction of time spent in the limiting states, $f_1 \in [0,1]$.

To derive Eq. 7.11, we start by expressing the variance as the difference between the expected value of the squared apparent FRET efficiency $\mathrm{E}\left[E_{\mathrm{app}}^2\right]$ and the square of the expected value $\mathrm{E}\left[E_{\mathrm{app}}\right]^2$, that is:

$$\mathrm{Var}\left(E_{\mathrm{app}}\right) = \mathrm{E}\left[E_{\mathrm{app}}^2\right] - \mathrm{E}\left[E_{\mathrm{app}}\right]^2. \tag{7.13}$$

Here, $\mathrm{E}[X^m]$ is the expected value of a continuous random variable X^m with $m \in \{1,2\}$, defined by:

$$\mathrm{E}[X^m] = \int X^m P(X)\, dX. \tag{7.14}$$

In BVA, the number of photons per sampling window, n, is constant. The probability to observe an apparent FRET efficiency, E_{app} , is thus equal to the probability of observing $n_A = nE_{\mathrm{app}}$ acceptor photons among the n detected photons, which is given by a binomial distribution:

$$P\left(E_{\mathrm{app}}\right) = P\left(n_A = nE_{\mathrm{app}} \middle| E_{\mathrm{app}}^{(i)}, n\right) = \binom{n}{n_A} {E_{\mathrm{app}}^{(i)}}^{n_A}\left(1-E_{\mathrm{app}}^{(i)}\right)^{n-n_A}, \tag{7.15}$$

where n_A is the number of acceptor photons and $E_{\mathrm{app}}^{(i)}$ is the ideal apparent FRET efficiency of the molecule in state *i*.

We assume that the conformational dynamics are slow compared to the sampling frequency of the standard deviation, which is defined by the average time needed to detect n photons. Then, in the presence of conformational dynamics between two states with apparent FRET efficiencies $E_{\mathrm{app}}^{(1)}$ and $E_{\mathrm{app}}^{(2)}$, the probability to observe a given average apparent FRET efficiency E_{app} in a sampling window of n photons is given by the weighted average of the binomial distributions for the two limiting states:

$$P\left(E_{\text{app}}\right) = f_1 P\left(n_A = n\,E_{\text{app}}\middle|E_{\text{app}}^{(1)}, n\right) + f_2 P\left(n_A = n\,E_{\text{app}}\middle|E_{\text{app}}^{(2)}, n\right), \tag{7.16}$$

where f_1 and f_2 are the probabilities that the molecule is in state 1 or 2, respectively (with $f_1 + f_2 = 1$). This follows because the molecule is found exclusively in one of the two limiting states during each sampling window of n photons.

We can now calculate the expected value of the apparent FRET efficiency E_{app} as:

$$\mathrm{E}\left[E_{\text{app}}\right] = f_1 E_{\text{app}}^{(1)} + f_2 E_{\text{app}}^{(2)}. \tag{7.17}$$

The expected value of the squared apparent FRET efficiency, $\mathrm{E}\left[E_{\text{app}}^2\right]$, is given by:

$$\begin{aligned}\mathrm{E}\left[E_{\text{app}}^2\right] &= \int E_{\text{app}}^2 P\left(E_{\text{app}}\right) dE_{\text{app}} \\ &= f_1 \int E_{\text{app}}^2 P\left(n_A = n\,E_{\text{app}}\middle|E_{\text{app}}^{(1)}, n\right) dE_{\text{app}} \\ &+ f_2 \int E_{\text{app}}^2 P\left(n_A = n\,E_{\text{app}}\middle|E_{\text{app}}^{(2)}, n\right) dE_{\text{app}}\end{aligned} \tag{7.18}$$

The integrals represent the expected value of the square, $\mathrm{E}[X^2]$, i.e., the second moments, for a binomial random variable. Using the definition of the variance as given in Eq. 7.13, $\mathrm{E}[X^2]$ can be calculated as:

$$\mathrm{E}[X^2] = \mathrm{Var}(X) + \mathrm{E}[X]^2 = np(1-p) + n^2p^2, \tag{7.19}$$

where n and p are the attempt number (i.e. number of photons in the averaging window, n) and success probability of the binomial process ($p = E_{\text{app}}^{(i)}$), and X is the number of successes (here, $X = n_A = nE_{\text{app}}$). The apparent FRET efficiency is given by $E_{\text{app}} = n_A/n$. Thus, for state i, the expected value of the squared apparent FRET efficiency, $\mathrm{E}^{(i)}\left[E_{\text{app}}^2\right]$, is given by:

$$\mathrm{E}^{(i)}\left[E_{\text{app}}^2\right] = \frac{E_{\text{app}}^{(i)}\left(1 - E_{\text{app}}^{(i)}\right)}{n} + E_{\text{app}}^{(i)}. \tag{7.20}$$

Finally, the variance of E_{app} is given by:

$$\begin{aligned}\mathrm{Var}^{(\mathrm{dyn})}\left(E_{\text{app}}\right) &= \mathrm{E}\left[E_{\text{app}}^2\right] - \mathrm{E}\left[E_{\text{app}}\right]^2 \\ &= f_1\left[\frac{E_{\text{app}}^{(1)}\left(1 - E_{\text{app}}^{(1)}\right)}{n} + {E_{\text{app}}^{(1)}}^2\right] + f_2\left[\frac{E_{\text{app}}^{(2)}\left(1 - E_{\text{app}}^{(2)}\right)}{n} + {E_{\text{app}}^{(2)}}^2\right] \\ &- \left(f_1 E_{\text{app}}^{(1)} + f_2 E_{\text{app}}^{(2)}\right)^2,\end{aligned} \tag{7.21}$$

where the fractions of the limiting states are defined by the observed average apparent FRET efficiency $E_{\text{app}} = f_1 E_{\text{app}}^{(1)} + f_2 E_{\text{app}}^{(2)}$. This is Eqn. 7.11. The fractions are given by:

$$f_1 = \frac{E_{\text{app}} - E_{\text{app}}^{(2)}}{E_{\text{app}}^{(1)} - E_{\text{app}}^{(2)}}; \quad f_2 = 1 - f_1. \tag{7.22}$$

The dynamic shift in BVA is then defined as the maximum difference between the static and dynamic FRET-lines:

$$ds_{(BVA)} = \max_{E_{\text{app}}} \left[\sigma_{E_{\text{app}}}^{(\text{dyn})}(E_{\text{app}}) - \sigma_{E_{\text{app}}}^{(\text{stat})}(E_{\text{app}}) \right]. \tag{7.23}$$

Since no analytical expression was found for Eq. 7.23, it was evaluated numerically.

Supplementary Note 8: R_0 determination for the Alexa546-Alexa647 dye pair for MalE

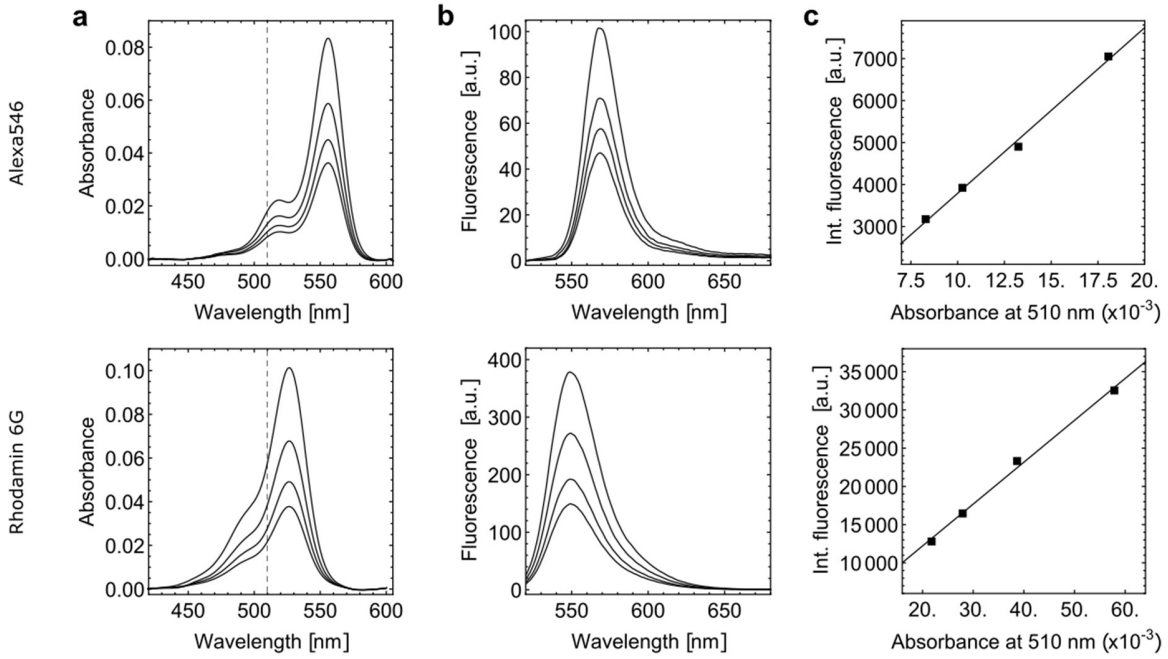
The Förster radius, R_0, is given by[15]

$$R_0{}^6 = \frac{9\,ln\,(10)}{128\,\pi^5\,N_A}\frac{\kappa^2}{n_{im}}Q_D\,\frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\,\lambda^4\,d\lambda}{\int_0^\infty F_D(\lambda)\,d\lambda}. \qquad (8.1)$$

where κ^2 is the orientation factor, Q_D the quantum yield of the donor, $F_D(\lambda)$ is the fluorescence emission spectrum of the donor, $\varepsilon_A(\lambda)$ is the absorption spectrum of the acceptor (scaled with the appropriate absorption coefficient), N_A is Avogadro's number and n_{im} is the index of refraction in the intervening medium. For an appropriate R_0 determination, it is necessary to determined the fluorescence quantum yield of the donor labeled to the molecule of interest as well as the overlap integral.

8.1 Fluorescence quantum yield $\Phi_{F,D}$

The fluorescence quantum yield of the donor dye Alexa 546 covalently bound to the protein was determined using Rhodamine 6G as a reference.



Supplementary Figure SN8.1. Exemplarily fluorescence quantum yield determination for MalE-1. (a) The absorption spectrum of Alexa546 (top) and Rhodamine 6G (bottom) at 0.5, 0.625, 0.75 and 1 µM concentration. **(b)** The emission spectrum of Alexa546 (top) and Rhodamine 6G (bottom) at 0.5, 0.625, 0.75 and 1 µM concentration excited at 510 nm. **(c)** The integrated fluorescence of Alexa546 (top) and Rhodamine 6G (bottom) plotted against absorbance at 510 nm at 0.5, 0.625, 0.75 and 1 µM concentration and linear fit to the data points (line). Assuming a fluorescence quantum yield of 91% for Rhodamine 6G, the quantum yield for Alexa546 was determined using the ratio of the slopes.

The absorption (Supplementary Figure SN8.1a) and emission spectra (Supplementary Figure SN8.1b) at 510 nm excitation wavelength were measured for four different dye concentrations in the range of 0.5-1.0 µM for the sub-stoichiometrically labeled donor only samples and compared to Rhodamine 6G in water. A linear fit to the integrated fluorescence intensity $\int_0^\infty F_D(\lambda)\, d\lambda$ for 510 nm excitation as a function of concentration provides two slopes: m_{Al546} and m_{R6G} (Supplementary Figure SN8.1c).

The fluorescence quantum yield of the donor Alexa546 is calculated from the slopes of the linear fit as

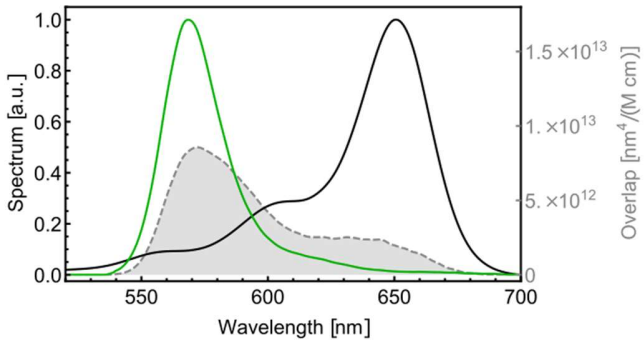
$$\boldsymbol{\Phi_{F,D}} = \frac{m_{Al546}}{m_{R6G}} \boldsymbol{\Phi_{F,R6G}} \tag{8.2}$$

where $\boldsymbol{\Phi_{F,R6G}} = 91 \pm 2\%$ is taken from literature[16]. The quantum efficiency of Alexa 546 bound to the protein was found to be $72 \pm 4\%$. This is similar to the quantum yield of Alexa546 alone in PBS given in the molecular probes handbook of 0.79[17].

8.2 Overlap integral J

The overlap integral was calculated using the emission spectrum $F_D(\lambda)$ of the donor only sample and a normalized absorption spectrum of the acceptor $\bar{\varepsilon}_A$ scaled to the literature extinction coefficient $\varepsilon_A(\lambda) = \varepsilon_{A_{max}}\, \bar{\varepsilon}_A$ using:

$$J = \frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\, \lambda^4\, d\lambda}{\int_0^\infty F_D(\lambda)\, d\lambda}. \tag{8.3}$$



Supplementary Figure SN8.2. The overlap integral of Alexa546-Alexa647. The normalized emission spectrum of Alexa546 (green) and absorption spectrum of Alexa 647 (black) plotted along with the overlap integral shown in gray for the MalE-1 mutant.

The overlap integral is illustrated in Supplementary Figure SN8.2 (gray area) resulting from the donor emission (green) and acceptor absorption (black) spectra.

8.3 Förster radius calculation

For calculating R_0, we have used the following values:

Orientation factor κ^2: 2/3

Averaged refractive index n_{im} (ref. 18): 1.4
Extinction coefficient at maximum $\varepsilon_{A_{max}}$ (ref. 18): 265,000 $OD/(M\ cm)$
Fluorescence quantum yield $\boldsymbol{\Phi_{F,D}}$: $72 \pm 4\%$
Overlap integral J: $7.0 \pm 0.1 \times 10^{15}\ nm^4/(M\ cm)$

and determined to be $R_0 = 65 \pm 3$ Å considering uncertainties in κ^2 of 10% and in n_{im} of 5%. This differs significantly from the manufacture's published R_0 of 74 Å[19].

The absorption and emission spectra were measured using singly labeled donor and acceptor mutants for MalE-1, MalE-2 and MalE-3. For other dye combinations used in this study, the following Förster radii were used: R_0 (Alexa546-Abberior STAR635P) = 62.5 Å, R_0 (Atto532-Atto643) = 59.0 Å and R_0 (Alexa488-Alexa647) = 52.0 Å.

Supplementary Note 9: Estimation of the residual anisotropies

To improve the robustness of the analysis, the residual anisotropies were determined via two approaches, from the steady-state intensities ("*ss*") and from the time-resolved data ("*tr"*). The values reported in Fig. 5e-f of the main text correspond to an average value of the two estimates.

a) Residual anisotropies from steady-state intensities

Using single-molecule multiparameter fluorescence detection (MFD)[5], steady-state anisotropies, r_{ss}, in each detected single-molecule event were calculated from the background-corrected photon counts registered in the parallel and perpendicular detection channels as described in detail in reference 20. Next, we plotted the obtained r_{ss} versus the intensity averaged fluorescence lifetime $\langle\tau\rangle_F$ to analyze their relation by the Perrin equation[21,22]. Here, we assume that the anisotropy decay can be described by two depolarization components[23]: (i) the global rotation of the protein, ρ_{glob}, of approx. 20 ns, and (ii) a second component of ~ 0.5 ns describing the local dye motions, ρ_{linker}, with amplitudes $x_{\text{global}} = 1 - x_{\text{linker}}$ and x_{linker}, respectively:

$$r_{ss} = r_0 \left(\frac{x_{\text{linker}}}{1 + \frac{\langle\tau\rangle_F}{\rho_{\text{linker}}}} + \frac{1 - x_{\text{linker}}}{1 + \frac{\langle\tau\rangle_F}{\rho_{\text{global}}}} \right). \tag{9.1}$$

The residual anisotropy is defined as the amplitude of the global rotational motion as:

$$r_{\infty,ss} = r_0(1 - x_{\text{linker}}), \tag{9.2}$$

where x_{linker} is given by:

$$x_{\text{linker}} = \left(\frac{r_{ss}}{r_0} - \frac{1}{1+\frac{\langle\tau\rangle_F}{\rho_{\text{global}}}}\right) \Big/ \left(\frac{1}{1+\frac{\langle\tau\rangle_F}{\rho_{\text{linker}}}} - \frac{1}{1+\frac{\langle\tau\rangle_F}{\rho_{\text{global}}}}\right), \tag{9.3}$$

and the fundamental anisotropy r_0 is assumed to be 0.38 for all dyes used here.

b) Residual anisotropies from the time-resolved anisotropy analysis

For the single-cysteine variants, the time-resolved anisotropy decay, $r(t)$, was obtained from ensemble TCSPC experiments. For some variants where single-cysteine mutants were not available, $r(t)$ was obtained from the single-molecule MFD experiments using a sub-ensemble selection of double-labeled molecules by applying an adequate value of the ALEX-2CDE filter. When needed, the bursts are additionally filtered using the stoichiometry value. From the selected double-labeled molecules, the fluorescence decays were obtained for the donor emission channel after donor excitation ($D_{ex}|D_{em}$) and the acceptor emission channel after acceptor excitation ($A_{ex}|A_{em}$). Note that, using this selection, the donor fluorescence lifetime is shortened by FRET to the acceptor. The anisotropy information was obtained from a global fit of the fluorescence decays $F_{VV}(t)$, $F_{VH}(t)$ and $F_{VM}(t)$, where VV and VH denote vertically (V) and horizontally (H) polarized emitted light upon vertical excitation (V), respectively and VM represents the fluorescence lifetime decay measured at the magic angle (~54.7˚) of the emission polarizer after vertical excitation. Here, the VM decay was not measured but rather constructed by mixing the VV and VH decays and with the experimentally determined g-factor and correcting for polarization mixing in the beam path of the high NA objective using the factors l_1 and l_2[20,24] as follows:

$$F_{VM}(t) = (1 - 3l_2)F_{VV}(t) + g_{VH/VV}(2 - 3l_1)F_{VH}(t) \tag{9.4}$$

By including the magic angle decay $F_{VM}(t)$ in a global analysis, the stability of the fit is significantly improved as it increases the robustness of the fluorescence lifetime estimation.

The experimental decay curves $F(t)$ are described based on the idealized model decays $f(t)$. The relation between time-resolved anisotropy and polarization resolved fluorescence decays $f(t)$ is given as follows[24] :

$$f_{VV}(t) = \frac{1}{3} f_{VM}(t)\big(1 + (2 - 3l_1)r(t)\big); \tag{9.5}$$

$$f_{VH}(t) = \frac{1}{3} f_{VM}(t)\big(1 - (1 - 3l_2)r(t)\big), \tag{9.6}$$

The magic angle decay, $f_{VM}(t)$, is modelled by a multi-exponential function:

$$f_{VM}(t) = F_0 \sum_i x_i \exp\left(-\frac{t}{\tau_i}\right), \tag{9.7}$$

where τ_i denotes the fluorescence lifetimes and x_i are the fraction of molecules with the corresponding lifetime, and F_0 is a scaling factor that normalizes the integral of the model decay

to the integrated experimental signal (i.e., the total measured counts). Similarly, the time-resolved anisotropy is modelled as:

$$r(t) = r_0\left(x_{\text{dye}}\exp\left(-\frac{t}{\rho_{\text{dye}}}\right) + x_{\text{linker}}\exp\left(-\frac{t}{\rho_{\text{linker}}}\right) + x_{\text{global}}\exp\left(-\frac{t}{\rho_{\text{global}}}\right)\right), \quad (9.8)$$

where ρ_j is the depolarization time of species j with anisotropy amplitude $b_j = r_0 x_j$. As above, we consider the global rotation of the molecule on timescales longer than 10 ns ($\rho_{\text{global}} > 10$ ns) and the local mobility of the tethered label including the linker on the nanosecond scale, which is accounted for using two components, one for the linker movement (ρ_{linker} ~ 1-5 ns) and the other for fast rotation of the dye (ρ_{dye} ~ 0.3-0.5 ns). The contribution of a specific process to the overall depolarization of the fluorescence signal is determined by the species fraction x_j, with $\sum_j x_j = 1$. The fundamental anisotropy is taken to be $r_0 = 0.38$ for all dyes. The contribution of global rotation of the macromolecule to the total depolarization of emitted light represents the residual anisotropy, $r_{\infty,tr}$, given by $r_{\infty,tr} = r_0 x_{\text{global}}$.

The model decays are then convolved with the corresponding polarized component of the instrument response function (IRF):

$$F_{VV}(t) = f_{VV}(t)\otimes IRF_{VV}(t) + sc_{VV}IRF_{VV}(t) + B_{VV}, \quad (9.9)$$

$$F_{VH}(t) = g_{VH/VV}\, f_{VH}(t)\otimes IRF_{VH}(t) + sc_{VH}IRF_{VH}(t) + B_{VH}, \quad (9.10)$$

$$F_{VM}(t) = f_{VM}(t)\otimes IRF_{VM}(t) + sc_{VM}IRF_{VM}(t) + B_{VM}, \quad (9.11)$$

Here, B_{VV}, B_{VH} and B_{VM} account for constant background signal, the factors sc_{VV}, sc_{VH} and sc_{VM} describe the contribution of scattered laser light, and the factor $g_{VH/VV}$ corrects for the different detection efficiencies in the two detection channels VV and VH and $\otimes$ denotes a circular convolution. Here we apply a circular convolution because the full microtime range is used for the analysis, resulting in a periodic signal. Note that the idealize curves, $f(t)$, have been normalized to the total measured number of counts in Eqs. 9.5 and 9.6 via Eq. 9.7. Prior to the convolution, the IRF is corrected for uncorrelated background signal, e.g., due to the detector dark counts, and a time shift of the IRF is applied to correct for shifts that arise due to count rate differences between the IRF and fluorescence decay measurements. Similarly to the $F_{VM}(t)$ decay, its corresponding $IRF_{VM}(t)$ was constructed by mixing $IRF_{VV}(t)$ and $IRF_{VH}(t)$ according to Eq. 9.4.

The quality of the fit was judged using the reduced chi-squared, χ_r^2. Analysis of the time-resolved anisotropy was done using the software package ChiSurf, a GUI-based suite composed of a collection of Python scripts for various types of fluorescence data analysis, available at https://github.com/Fluorescence-Tools/chisurf.

c) Computation of the average residual and combined anisotropies from steady-state and time-resolved measurements

The residual anisotropies from the two approaches were averaged to compute the average residual anisotropy, $\langle r_\infty \rangle_{tr,ss}$:

$$\langle r_\infty \rangle_{tr,ss} = \frac{r_{\infty,tr} + r_{\infty,ss}}{2} \tag{9.12}$$

The uncertainty $\Delta\langle r_\infty \rangle_{tr,ss}$ was obtained as the standard deviation of the two estimates, corresponding to:

$$\Delta\langle r_\infty \rangle_{tr,ss} = \frac{\left| r_{\infty,tr} - r_{\infty,ss} \right|}{\sqrt{2}} \tag{9.13}$$

The combined residual anisotropy was then computed as a geometric average:

$$\langle r_{c,\infty} \rangle_{tr,ss} = \sqrt{\langle r_{\infty,D} \rangle_{tr,ss} \cdot \langle r_{\infty,A} \rangle_{tr,ss}}, \tag{9.14}$$

and the corresponding error was obtained using standard error propagation as:

$$\Delta\left(\langle r_{c,\infty} \rangle_{tr,ss}\right) = \frac{1}{2}\sqrt{\frac{\langle r_{\infty,D} \rangle_{tr,ss}}{\langle r_{\infty,A} \rangle_{tr,ss}} (\Delta\langle r_{\infty,A} \rangle_{tr,ss})^2 + \frac{\langle r_{\infty,A} \rangle_{tr,ss}}{\langle r_{\infty,D} \rangle_{tr,ss}} (\Delta\langle r_{\infty,D} \rangle_{tr,ss})^2}. \tag{9.15}$$

Supplementary Note 10: Calculation of distance uncertainties according to the "diffusion with traps" model

Anisotropy measurements of fluorophores tethered to the surface of biomolecules indicate that the dye motion is commonly not isotropic despite the use of long, flexible linkers. Estimating a precise value of the orientational factor κ^2 is generally impossible as it would require knowledge of the mutual orientations of the transition dipole moments of the donor and acceptor molecules and their orientation with respect to the inter-dye distance vector. However, one can restrict the possible κ^2 values to a range that is compatible with the experiments with the help of the experimental anisotropy data. Previous approaches have estimated the uncertainty in κ^2 from the residual anisotropy in terms of rotational restrictions[25,26]. Here, we adopt a different approach termed the *diffusion with traps* (DWT) model[27] that provides an uncertainty in the measured donor-acceptor separation due to the unknown orientation factor, κ^2, based on the experimentally obtained fractions of trapped dye species and FRET efficiencies. The fraction of trapped dye species is obtained from the fluorescence anisotropy decays of the donor and acceptor fluorophores as a ratio of the residual, r_∞, to the fundamental anisotropy, r_0, which is often referred to as the second rank order parameter, $S^{(2)}_{A/D}$:

$$S^{(2)}_{A/D} = \frac{r_{\infty,A/D}}{r_{0,A/D}}, \tag{10.1}$$

where the subscript denotes that the order parameter is computed either for the donor (D) or acceptor (A) fluorophore.

Inspired by observations from MD simulations, the DWT model assumes that the fluorophores are either completely free, i.e., undergoing isotropic reorientation, or remain immobile for an extended period of time, meaning that they are trapped. It is also assumed that the exchange between these two species is slower than the donor fluorescence lifetime (i.e., static on the timescale of FRET)

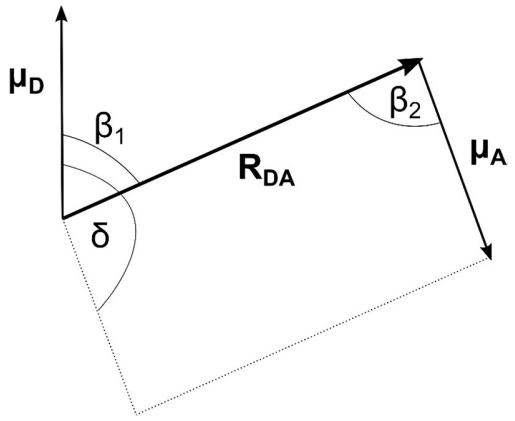
but faster than the diffusion time through the confocal volume in smFRET experiments of ~1-5 ms. This approximation is supported from MD data where the residence times of these states are found to be on the order of 10-100 ns or longer.
The DWT model assumes a mixture of trapped and free dye species and considers all four mobility scenarios for the donor and acceptor when computing the FRET efficiency, that is:

(I) $D_{\text{free}} - A_{\text{free}}$
(II) $D_{\text{free}} - A_{\text{trapped}}$
(III) $D_{\text{trapped}} - A_{\text{free}}$ and
(IV) $D_{\text{trapped}} - A_{\text{trapped}}$.

To calculate the range of possible κ^2 values, we generate random orientations of the dipole moment vectors, $\boldsymbol{\mu_D}$ and $\boldsymbol{\mu_A}$. For a given orientation of dipole vectors, the orientation factor κ^2 is given by[27]:

$$\kappa^2 = \frac{2}{3} + \frac{2}{3}S_D^{(2)}S_{\beta_1}^{(2)} + \frac{2}{3}S_A^{(2)}S_{\beta_2}^{(2)} + \frac{2}{3}S_D^{(2)}S_A^{(2)}\left[S_\delta^{(2)} + 6S_{\beta_1}^{(2)}S_{\beta_2}^{(2)} + 1 + 2S_{\beta_1}^{(2)} + 2S_{\beta_2}^{(2)} - 9\cos\beta_1\cos\beta_2\cos\delta\right] \quad (10.2)$$

where the angles β_1 and β_2 define the dipole orientations with respect to the donor-acceptor distance vector $\mathbf{R_{DA}}$ for D and A, respectively, and δ corresponds to the angle between the D and A transition dipole vectors, as illustrated below:



The second-rank order parameters for the angles are defined as follows:

$$S_\delta^{(2)} = \frac{1}{2}(3cos^2\delta - 1) \quad (10.3)$$

$$S_{\beta_1}^{(2)} = \frac{1}{2}(3cos^2\beta_1 - 1) \quad (10.4)$$

$$S_{\beta_2}^{(2)} = \frac{1}{2}(3cos^2\beta_2 - 1) \tag{10.5}$$

In the first step, generated are *N* random vectors, which represent random orientations of transition dipole vectors of donor and acceptor dyes. From those vectors, the angles β_1, β_2 and δ are calculated. Since we do not use the FRET sensitized acceptor anisotropy decay, δ is allowed to vary between 0° and 90°. However, it is possible to limit the range of possible δ - and κ^2- values by using the residual FRET sensitized acceptor anisotropy, $r_{\infty,A(D)}$.

In the next step, the FRET efficiency is computed as a function of κ^2 by considering all possible scenarios (I - IV) weighted by their respective probability of occurrence based on the estimated fractions of free/trapped dyes from the order parameters:

$$\begin{aligned} E(\kappa^2) = {} & \left(1 - S_D^{(2)}\right)\left(1 - S_A^{(2)}\right)E + S_D^{(2)}S_A^{(2)}E\left(\kappa^2(D_{\text{trapped}} - A_{\text{trapped}})\right) \\ & + S_D^{(2)}\left(1 - S_A^{(2)}\right)E\left(\kappa^2(D_{\text{trapped}} - A_{\text{free}})\right) \\ & + \left(1 - S_D^{(2)}\right)S_A^{(2)}E\left(\kappa^2(D_{\text{free}} - A_{\text{trapped}})\right) \end{aligned} \tag{10.6}$$

where $E = E(\kappa^2(D_{\text{free}} - A_{\text{free}})$ is the experimentally measured FRET efficiency, and:

$$\begin{aligned} & E\left(\kappa^2(D_{\text{trapped/free}} - A_{\text{trapped/free}})\right) \\ & \qquad = \frac{1}{1 + \frac{2/3}{\kappa^2(D_{\text{trapped/free}} - A_{\text{trapped/free}})}\left(\frac{1}{E} - 1\right)}, \end{aligned} \tag{10.7}$$

where the factor $2/3$ represents the isotropically averaged κ^2 value. The value of $\kappa^2(D_{\text{trapped/free}} - A_{\text{trapped/free}})$ is calculated according to Eq. 10.2 using the following assignment of the order parameters:

$$\begin{aligned} & \kappa^2\left(D_{\text{trapped}} - A_{\text{trapped}}\right)\text{: } S_D^{(2)} = S_A^{(2)} = 1 \\ & \kappa^2\left(D_{\text{trapped}} - A_{\text{free}}\right)\text{: } S_D^{(2)} = 1, \qquad S_A^{(2)} = 0 \\ & \kappa^2\left(D_{\text{free}} - A_{\text{trapped}}\right)\text{: } S_D^{(2)} = 0, \qquad S_A^{(2)} = 1 \end{aligned} \tag{10.8}$$

The dynamically averaged $\langle\kappa^2\rangle$ that corresponds to the obtained FRET efficiency $E(\kappa^2)$ is then obtained as:

$$\langle\kappa^2\rangle = \frac{2}{3}\frac{\frac{1}{E} - 1}{\frac{1}{E(\kappa^2)} - 1}, \tag{10.9}$$

where again the factor $2/3$ represents the isotropic average of κ^2.

This procedure is repeated for 10,000 randomly generated orientations of the dye vectors for a pre-defined distance, $R_{DA}(2/3) = R_0(1/E - 1)^{\frac{1}{6}}$, where E is the measured FRET efficiency and R_0 is

the Förster radius assuming a κ^2 of 2/3[rds]. The obtained dynamically-averaged $\langle\kappa^2\rangle$ values are then converted into a normalized distance R_{app} according to the following equation:

$$R_{\mathrm{app}} = \frac{R_{DA}(\langle\kappa^2\rangle)}{R_{DA}(2/3)} = \left(\frac{3}{2}\langle\kappa^2\rangle\right)^{-1/6} \quad (10.10)$$

The relative standard error in distance due to κ^2, $\Delta R_{\mathrm{app}}(\kappa^2)$, is then defined as the standard deviation of the obtained R_{app} distribution normalized to its mean (Fig. 5e).

Supplementary Note 11: Donor quenching estimation at different labeling positions on the MalE protein.

To explain the observed quenching at particular locations in MalE, we used coarse-grained Brownian dynamics simulations as described previously[1]. First, we identified the positions of the amino acids that tend to quench. These are typically MET, TYR, TRP and HIS. Next, we assigned is a quenching rate constant to each of them. We found that the experimental data is explained best with a quenching rate constant of k_Q = 2 ns^{-1}. After the position and quenching rate of the problematic amino acids are identified, we determine all sterically allowed positions of the donor dye using Accessible Volume simulations[23]. Within the accessible volume, we simulate diffusion of a dye using Brownian dynamics. As it is known that dyes diffuse slower in the vicinity of biomolecule's surface due to the non-specific sticking interactions, a heterogeneous diffusion model was applied. In this model, the diffusion coefficient of D = 10 Å^2 ns^{-1} is decreased by a factor of 10 in the vicinity of the surface. For each dye position within the AV, we measured the distance between the dye and all C_β atoms of the protein. Whenever the minimum distance was below the threshold of $R_{surface}$ = 8 Å, the slow diffusion coefficient (D = 1 Å^2 ns^{-1}) was used. Next, for each dye position, we estimated the fluorescence properties by measuring the distance of the fluorophore to the amino acids that can quench. The quenching was approximated by a step function such that whenever the dye was found within a distance of R_{rad} = 8.5 Å, quenching occurs at the given rate of k_Q. Hence, during quenching, the emission rate of the donor is given by $k_D = \tau_0^{-1} + k_Q$ and, for distances > R_{rad}, the donor is unquenched $k_D = \tau_0^{-1}$. Simulations were performed using the open-source GUI version of *QuEST – Quenching Estimator* software (GitHub page: https://github.com/Fluorescence-Tools/quest). Parameters used in the simulations are summarized in the Supplementary Table SN11.1 and our estimated fluorescence quantum yields for the donor at different labeling positions of MalE protein are provided in Supplementary Table SN11.2.

Supplementary Table SN11.1. Parameters used for the coarse-grained MD simulations of Alexa546 quenching at the different labeling positions of MalE in the apo and holo states.

Dye species	**τ_0 / ns**	**Dye properties/ Å**		**D_{free} / Å^2 ns^{-1}**	**$R_{surface}$ / Å**	**$D_{surface}$ / Å^2 ns^{-1}**	**R_{rad} / Å**	**Quencher**	**k_Q /ns^{-1}**	**BD simulations**	
Alexa Fluor 546	4.1	L_{length}	20.5	10	8	1	8.5	MET	2	Simulation time / µs	10
								TYR		Time step / ps	2
		L_{width}	4.5					TRP		Grid size / Å	0.5
		R_{dye}	3.5					HIS			

Supplementary Table SN11.2. Simulated fluorescence quantum yield $\phi_{F,D}$ and fraction of dye states that collided with quenching amino acids for Alexa546 dye on the MalE protein. Using the coarse-grained Brownian dynamics simulations (see used simulation parameters in Supplementary Table SN11.1), we estimated quenching of the donor dye at the labeling positions of MalE. The positions that show a prominent quenching in BD simulations are S352C (apo/holo), K34C(apo/holo) and T36C (apo/holo), as given by frequent collisions with quenching amino acids. Using species averaged fluorescence lifetime $\langle\tau\rangle_x$ obtained by fitting ensemble time-resolved lifetime measurements of single labelled cysteine mutants (Supplementary Table 9), one can determine experimental fluorescence quantum yields, according to $\phi_{F,D} = \phi_{F,D,ref}\frac{\langle\tau_{D(0)}\rangle_x}{\tau_{D(0)}}$, with $\phi_{F,D,ref}$ = 0.72 (see Supplementary Note 8) and $\tau_{D(0)}$ = 4.1 ns. As predicted by BD simulations, experiments confirm that at position S352C dye is prone to more frequent collisions with quenchers, resulting in lower fluorescence quantum yield compared to other labelling positions.

Position	State	Fraction of collisions with quenchers / %	Simulated $\phi_{F,D}$	Experimental $\phi_{F,D}$*
K29C	apo	10.84	0.64	0.71
	holo	14.45	0.60	0.71
S352C	apo	54.77	0.50	0.66
	holo	42.73	0.52	0.67
D87C	apo	23.16	0.57	0.68
	holo	28.65	0.53	0.69
A186C	apo	20.95	0.62	0.69
	holo	20.94	0.61	0.69
A134C	apo	7.92	0.64	0.69
	holo	5.96	0.65	0.70
K34C	apo	38.62	0.58	-
	holo	32.88	0.58	-
T36C	apo	49.94	0.50	-
	holo	48.02	0.53	-
N205C	apo	28.94	0.52	-
	holo	25.08	0.55	-

*Ensemble lifetime measurements of the single-labeled cysteine mutants K34C, T36C and N205C are not available.

Supplementary Note 12: Determination of statistical significance of the dynamic shifts of MalE.

To assess the statistical significance of the excess dynamic shift obtained for the measured MalE mutants, we compared the distribution of the measured dynamic shift between the different labs to the dynamic shifts observed for dsDNA. We computed the *p*-value for each of them using a chi-square score. With the *p*-value, we test whether distributions of ds values differ between DNA samples (reference, *r*) and the measured MalE and U2AF65 samples (*s*). Our null hypothesis is that measured MalE and U2AF65 samples appear as static as DNA. The *p*-value is the probability that the data is observed given the null hypothesis and can be used to decide whether the null hypothesis is valid or should be rejected given. Small *p*-values, particularly those below 0.05 (corresponding to a 2σ confidence interval), leads to the rejection of the null hypothesis.

Assuming that the reference DNA samples and samples under the tested hypothesis (MalE and U2AF65 variants) both have normally distributed *ds* values with mean $\langle ds \rangle_r$, $\langle ds \rangle_s$ and standard error of the mean SEM_r, SEM_s respectively, the chi-square test is then performed as follows:

$$\chi^2 = \left(\frac{\langle ds \rangle_r - \langle ds \rangle_{tot}}{SEM_r} \right)^2 + \left(\frac{\langle ds \rangle_s - \langle ds \rangle_{tot}}{SEM_s} \right)^2 \tag{12.1}$$

with $\langle ds \rangle_{tot}$ being:

$$\langle ds \rangle_{tot} = \frac{\frac{\langle ds \rangle_r}{{SEM_r}^2} + \frac{\langle ds \rangle_s}{{SEM_s}^2}}{\frac{1}{{SEM_r}^2} + \frac{1}{{SEM_s}^2}} \tag{12.2}$$

Using the probability density function of χ^2 distribution:

$$f(\chi^2 | N_{dof}) = \frac{1}{2^{\frac{N_{dof}}{2}} \Gamma\left(\frac{N_{dof}}{2}\right)} (\chi^2)^{\frac{N_{dof}}{2}-1} e^{-\frac{\chi^2}{2}} \tag{12.3}$$

where N_{dof} is number of degrees of freedom (N_{dof} =$N_{measurements} - N_{fit.params}$; here, $N_{dof} = 1$), one can determine the *p*-value, also referred to as *significance* or *certainty* that a reference distribution (DNA) does not match the sampled one (MalE, U2AF65). The p-value quantifies how unlikely it is to obtain a χ^2 value that is larger than the one observed between reference and measured sample. Commonly, the *p*-value is defined as the area under the right tail of the $f(\chi^2 | N_{dof})$ function:

$$p = \int_{\chi^2}^{+\infty} f(\chi^2 | N_{dof}) d\,\chi^2 \tag{12.4}$$

Using the described methodology in Supplementary Note 6, three labs estimated the apparent dynamic shift, *ds*, for different dye combinations of MalE and U2AF65 mutants (Supplementary Table 7). Additionally, one lab determined *ds* values for the DNA rulers, that would serve as a reference value of what would be the observed *ds* value for static systems. Furthermore, using *ds* values reported from all three labs, we calculated *p*-values for all measured variants (Supplementary Table SN12.1). *p*-values below 0.05 are obtained for the apo state of MalE-1, MalE-5 and U2AF65, and for the holo state of MalE-1, MalE-2, MalE-4, MalE-5 and U2AF65. For those samples, the distribution of *ds* values cannot be explained with the one of reference DNA sample and the samples appear to be more dynamic compared to dsDNA.

Furthermore, *p*-values were computed using the residual dynamic shift values after filtering out dyes with pronounced sticking interactions. Significant shrinking in average *ds* value was obtained mainly for MalE-1, where the *p*-value could not be estimated due to only one point being left after dye filtering. An increase in the *p*-value can be observed for a few mutants, meaning that, after removal of sticking artifacts, they do not appear significantly different compared to a static DNA sample. This effect is particularly pronounced for the MalE-4 apo/holo sample.

Supplementary Table SN12.1. *p*-values for MalE and U2AF65 samples before and after filtering dye-pairs with pronounced sticking interactions. *p*-values were computed according to Eq. 12.4.

	***p*-value all dyes**		***p*-value dyes ($\Delta R_{app}(\kappa^2) \leq 10\%$)**	
sample \ **condition**	**apo**	**holo**	**apo**	**holo**
MalE-1	0.0004	0.0024	n.a*	n.a*
MalE-2	0.7841	0.0376	0.7841	0.0376
MalE-3	0.9290	0.8973	0.5959	0.8973
MalE-4	0.0679	0.0051	0.4009	0.2024
MalE-5	1.20e-6	1.38e-6	0.0002	1.08e-5
U2AF65	1.11e-52	0.0009	6.91e-36	0.0260

**p*-value could not be estimated due to only one point being left after filtering dyes with pronounced sticking interactions

Supplementary Note 13: Estimation of conformational flexibility from the residual dynamic shift for MalE

The dynamic shift of a population in the E-τ plot is defined as the minimum distance to the reference static FRET-line given by:

$$E = 1 - \frac{\langle \tau_{D(A)} \rangle_F}{\tau_{D(0)}}, \tag{13.1}$$

where $\langle \tau_{D(A)} \rangle_F$ is the intensity-weighted average donor fluorescence lifetime and $\tau_{D(0)}$ is the donor-only lifetime (see Supplementary Note 6). We use the normalized donor fluorescence lifetime $\langle \tau_{D(A)} \rangle_F / \tau_{D(0)}$ to quantify the dynamic shift as this ensures that both axes in the E-τ plot range from 0 to 1. We have previously derived an expression for the maximum dynamic shift of a two-state system from the static FRET-line in a plot of E against $\langle \tau_{D(A)} \rangle_F / \tau_{D(0)}$ as a function of the limiting FRET efficiencies of the two states, E_1 and E_2 (Eq. 7.1 and ref. 14):

$$\mathrm{ds} = \frac{1}{\sqrt{2}} \left(\sqrt{1 - E_1} - \sqrt{1 - E_2} \right)^2 \tag{13.2}$$

Here, we use Eq. 13.2 to estimate the conformational flexibility from the measured dynamic shift under the following assumptions.

1) The dynamics are fast compared to the burst duration (< 100 µs) so that complete averaging occurs.

2) The two states are equally populated, that is the equilibrium constant of the dynamics, $K = k_{12}/k_{21}$, is equal to 1.

3) The dynamics are symmetric around the FRET-averaged mean interdye distance $R_{\langle E \rangle}$, i.e. the two states have interdye distances of $R_{\langle E \rangle} \pm \delta R$.

Under these assumptions, the dynamic population will be at a defined position off the static line that falls between the two limiting states. The dynamic shift then only depends on the amplitude of the distance fluctuation, δR:

$$\mathrm{ds}(\delta R) = \frac{1}{\sqrt{2}} \left(\sqrt{1 + \left(\frac{R_0}{R_{\langle E \rangle} - \delta R} \right)^6} - \sqrt{+ \left(\frac{R_0}{R_{\langle E \rangle} + \delta R} \right)^6} \right)^2 . \tag{13.3}$$

We numerically solve this equation to translate the measured dynamic shift of a given dye pair into an apparent distance fluctuation given the Förster radius R_0 and FRET-averaged interdye distance $R_{\langle E \rangle}$. The estimated distance fluctuations δR are given in Supplementary Table 7.

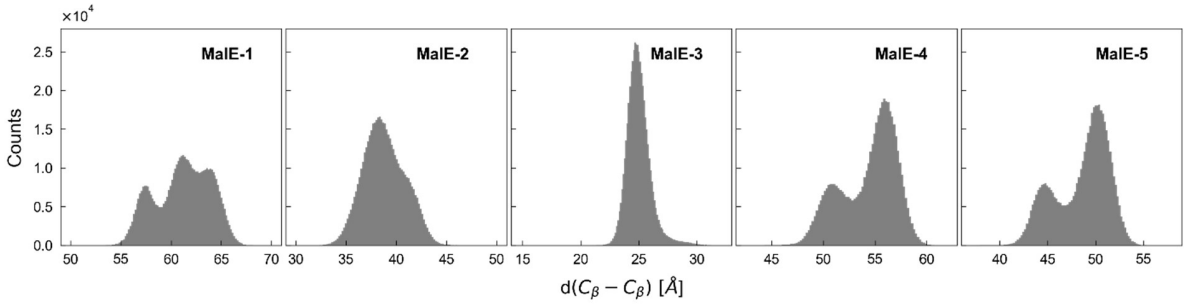
Supplementary Note 14: Molecular dynamics (MD) simulations of MalE.

Setup of topology and coordinate files. All-atom MD simulations were performed for MalE using the Amber18 suite[28] of programs with the FF14SB protein force-field[29] and the TIP4P-Ew water model[30]. For the initial structure, the crystal structure from PDB ID 1OMP was used. Using the LEaP program, the initial structure was solvated using TIP4P-Ew water molecules in an octahedral box such that the distance between the edge of water box and closest solute atom was at least 11Å. K^+/Cl^- ions were added to neutralize the system and then additional ions were added to reach a concentration of 50 mM for both ions. To allow for a larger time step of 4 fs, Hydrogen Mass Repartitioning (HMR)[31] was applied using ParmEd in the Amber software suite.

Minimization and Thermalization. After the topology and coordinate files were created and solvated, energy minimization of the system was performed while the positional restraints were applied to the solute atoms. Minimization was performed in two phases: first with high and then with low force constants for the positional restraints. In the first phase, the positional restraint was applied using a harmonic potential with a force constant of 25 kcal mol^{-1} $Å^{-2}$ and minimization was done through a total of 15,000 steps (5,000 steps of steepest descent minimization followed by 10,000 steps of conjugate gradient minimization). In the second phase, the force constant was reduced to 5 kcal mol^{-1} $Å^{-2}$ and minimization was done in the same manner as the first phase with 5,000 steps of steepest descent minimization followed by 10,000 steps of conjugate gradient minimization. After minimization, the system was heated from 100-300 K over 100 ps using NVT-MD simulations with the force constant of 5 kcal mol^{-1} $Å^{-2}$. The temperature change was applied in steps as follows. In the first 80 ps, the temperature was gradually raised from 100 K to the target 300 K. In the last 20 ps, the system was kept at the target temperature. To adjust the solvent density after increasing the temperature, heating was followed with a 300 ps of MD simulations using the NPT ensemble and the same force constant. Lastly, the force constant of the harmonic restraints was gradually reduced from 5 kcal mol^{-1} $Å^{-2}$ to 1 kcal mol^{-1} $Å^{-2}$ over 5×20 ps of NVT-MD simulations and then unrestrained NVT-MD production runs were performed.

Production runs. In total, 5 independent production replicas with a length of 2 μs each were performed. Conformations were saved every 20 ps. The SHAKE algorithm[32] was used in the production runs to constrain the bond lengths involving hydrogen atoms and long-range electrostatic interaction were treated using the Particle Mesh Ewald method[33] with a non-bonded cutoff of 8 Å.

Trajectories were analyzed using the MDTraj[34] library. The distances between the C_β atoms of the labeling positions for the five MalE mutants were calculated. For mutants MalE-1, MalE-4 and MalE-5, larger C_β-C_β distance fluctuations were observed.



Supplementary Figure SN14.1. Distributions of C_β-C_β distances from MD simulations of MalE proteins in the apo state. For each of the MalE mutants in the absence of fluorescent labels, the distance distribution between C_β atoms of the labeling positions is shown from a total of five production runs of 2µs length each. Consistent with smFRET experiments, a higher backbone-to-backbone distance fluctuations are detected for mutants MalE-1, MalE-4 and MalE-5. The obtained mean and standard deviations are $\langle d(C_\beta - C_\beta)\rangle_{MalE-1}$ = (61.2 ± 2.7) Å, $\langle d(C_\beta - C_\beta)\rangle_{MalE-2}$ = (38.8 ± 2.1) Å, $\langle d(C_\beta - C_\beta)\rangle_{MalE-3}$ = (25.0 ± 1.0) Å, $\langle d(C_\beta - C_\beta)\rangle_{MalE}$ = (54.4 ± 2.7) Å and $\langle d(C_\beta - C_\beta)\rangle_{MalE-5}$ = (48.6 ± 2.8) Å.

Supplementary Note 15: Theoretical limits for detecting dynamics

The theoretical description of the dynamic shifts in the E-τ and BVA plots as given in Supplementary Note 6 allows us to assess the detection limit for conformational dynamics with respect to the uncertainty of the experiment. To define the detectability of a given shift from the static FRET-line in an experiment, it is necessary to quantify the experimental uncertainties associated with the FRET indicators E and $\langle\tau_{D(A)}\rangle_F$ for the E-τ plot, and E_{app} and $\sigma_{E_{\text{app}}}^{(\text{stat})}$ for the BVA plot. The experimental uncertainty contains contributions from the photon counting statistics of the experiment and systematic errors due to the calibrations or timescales of the dynamic exchange.

We first consider only the statistical error based on the standard error of the mean (σ_{SEM}) of a population, given by:

$$\sigma_{SEM} = \frac{\sigma_{SN}}{\sqrt{N_b}}, \tag{15.1}$$

where σ_{SN} is the shot-noise broadened width of a population and N_b is the number of single molecule bursts in the population. A given dynamic shift is considered detectable if it is larger than the SEM of a population on the static FRET-line.

I. Detection limit for dynamics in the E-τ plot

Shot-noise broadening in the E-τ plot

To simplify the expressions for the shot-noise broadening of a population in the E-τ plot, we assume that all detected single-molecule events have the same number of photons, N_p. For the FRET efficiency, the population width is then given by[35]:

$$\sigma_{SN}^{E} = \sqrt{\frac{E(1-E)}{N_p}}. \tag{15.2}$$

For the donor fluorescence lifetime $\tau_{D(A)}$, the variance of the estimator in the absence of background and for an instrument response with zero width (δ-function) is (ref 36,37):

$$\text{Var}\left(\tau_{D(A)}\right) = \frac{1}{N_p}{\tau_{D(A)}}^2 \frac{k^2}{r^2}(1-e^{-r})\left(\frac{e^{\frac{r}{k}}(1-e^{-r})}{\left(e^{\frac{r}{k}}-1\right)^2} - \frac{k^2}{e^r-1}\right)^{-1}, \tag{15.3}$$

where k is the number of TCSPC detection channels and $r = T/\tau_{D(A)}$ is the ratio of the detection time window T and the fluorescence lifetime $\tau_{D(A)}$. The complex expression for the variance arises

because only a part of the decay is seen in the finite time window T and due to the discretization of the time axis into k bins. In the limit of a large time window T and a high number of bins k, the expression simplifies to:

$$\lim_{r\to\infty,\ k\to\infty} \mathrm{Var}\left(\tau_{D(A)}\right) = \frac{1}{N_p}{\tau_{D(A)}}^2 \tag{15.4}$$

The same result is obtained by considering that the sum of independent exponential random variables δt_i (i.e. delay times) follows the Erlang distribution with variance:

$$\mathrm{Var}\left(\sum_{i=1}^{N_p} \delta t_i\right) = N_p {\tau_{D(A)}}^2 \tag{15.5}$$

from which the variance of the lifetime estimate is obtained as:

$$\mathrm{Var}\left(\frac{1}{N_p}\sum_{i=1}^{N_p} \delta t_i\right) = \frac{1}{N_p^2}\mathrm{Var}\left(\sum_{i=1}^{N_p} \delta t_i\right) = \frac{1}{N_p}{\tau_{D(A)}}^2 \tag{15.6}$$

Thus, for the normalized donor lifetime, we obtain the shot-noise limited width of the burst-wise distribution as:

$$\sigma_{SN}^{\tau} = \frac{\tau_{D(A)}}{\sqrt{N_p}\ \tau_{D(0)}} \tag{15.7}$$

To obtain the width of the distribution along the vector of the dynamic shift, i.e., orthogonal to the static FRET-line, the two contributions are combined geometrically:

$$\sigma_{SN}^{(E-\tau)} = \frac{1}{\sqrt{2}}\sqrt{{\sigma_{SN}^{E}}^2 + {\sigma_{SN}^{\tau}}^2}, \tag{15.8}$$

The standard error of the mean of the population, determined for example by Gaussian fitting, depends on the number of single molecule events in the population N_b, and is given by:

$$\sigma_{SEM}^{(E-\tau)} = \frac{\sigma_{SN}^{(E-\tau)}}{\sqrt{N_b}} \tag{15.9}$$

Effect of calibration error in the γ-factor

An incorrect calibration of the detection efficiency correction factor γ can lead to the detection of false-positive apparent dynamic shifts by shifting the population away from the static FRET line. However, only an underestimation of the γ-factor, leading to an overestimation of the FRET efficiency E, will shift the population above the static FRET-line, whereas an overestimation of γ will result in an unphysical shift below the static FRET-line. The propagated uncertainty of the γ-factor on the FRET efficiency is derived in Supplementary Note 5 (Eq. 5.5) and is given by:

$$\sigma_{E,\gamma} = E(1-E)\frac{\Delta\gamma}{\gamma}. \tag{15.10}$$

The calibration error, $\sigma_{E,\gamma}$, is combined with the shot-noise related uncertainty $\sigma_{SN}^{(E-\tau)}$ as:

$$\sigma_{SN,\gamma}^{(E-\tau)} = \sqrt{\sigma_{SN}^{(E-\tau)^2} + \sigma_{E,\gamma}^2}, \tag{15.11}$$

leading to an estimation of the combined standard error of the mean of:

$$\sigma_{SEM}^{(E-\tau)} = \frac{\sigma_{SN,\gamma}^{(E-\tau)}}{\sqrt{N_b}}. \tag{15.12}$$

II. Detection limit for dynamics in the BVA plot

Shot-noise broadening in the BVA plot

For BVA, we only have to consider the broadening along the y-axis, i.e., the uncertainty in the variance estimate. Assuming the apparent FRET efficiency, E_{app}, follows a Gaussian distribution with width parameter $\sigma_{E_{\mathrm{app}}}$, the estimated standard deviation in BVA, $\tilde{\sigma}_{E_{\mathrm{app}}}$, follows a scaled chi-squared distribution:

$$(M-1)\frac{\tilde{\sigma}_{E_{\mathrm{app}}}^2}{\sigma_{E_{\mathrm{app}}}^2} \sim \chi_{M-1}^2. \tag{15.13}$$

Here, $\sigma_{E_{\mathrm{app}}} = \sqrt{\frac{E_{\mathrm{app}}(1-E_{\mathrm{app}})}{n}}$ and M is the number of samples for the standard deviation estimate given by $M = N_p/n$, where N_p is the number of photons and n is the photon averaging window used for BVA. The upper 1σ confidence interval was used for the standard deviation estimate $\tilde{\sigma}_{E_{\mathrm{app}}}$ and is given by:

$$P\left((M-1)\frac{\tilde{\sigma}_{E_{\mathrm{app}}}^2}{\sigma_{E_{\mathrm{app}}}^2} \le (M-1)\frac{\tilde{\sigma}_{E_{\mathrm{app,UL}}}^2}{\sigma_{E_{\mathrm{app}}}^2}\right) = T_M(1), \tag{15.14}$$

where $\tilde{\sigma}_{E_{\mathrm{app,UL}}}^2$ is the 1σ upper limit and $T_M(x)$ is the cumulative Student's t-distribution with M degrees of freedom. Assuming an average of N_p = 100 photons per burst and a photon averaging window of n = 5, then M = 20 and $T_{20}(1) \approx 0.84$. The upper limit $\tilde{\sigma}_{E_{\mathrm{app,UL}}}$ is given by:

$$\tilde{\sigma}_{E_{\mathrm{app,UL}}} = \sqrt{\frac{E_{\mathrm{app}}(1-E_{\mathrm{app}})}{n}}\sqrt{\frac{\chi_{\mathrm{inv},M-1}^2\left(T_M(1)\right)}{M-1}}, \tag{15.15}$$

where $\chi^2_{\mathrm{inv},M-1}(x)$ is the inverse cumulative chi-squared distribution with $M-1$ degrees of freedom.

The shot-noise broadening around the true value on the static FRET-line is then given by:

$$\sigma_{SN}^{(BVA)} = \tilde{\sigma}_{E_{\mathrm{app}},\mathrm{UL}} - \sqrt{\frac{E_{\mathrm{app}}\left(1-E_{\mathrm{app}}\right)}{N}}, \tag{15.16}$$

from which the standard error of the mean is obtained by dividing by the square root of the number of bursts (Eqn. 15.9).

Kinetic averaging in BVA

As BVA relies on the sampling of the FRET efficiency distribution from the photon time trace, it is subject to kinetic averaging during the sampling time, reducing the detectable variance. Thus, the dynamic shift is decreased for dynamics that occurs on the order of the sampling time or faster.

The sampling time, T, depends on the observed signal count rate, S, and the photon number, n, used to sample the FRET efficiency. The delay time between subsequent photon detection events is exponentially distributed. The time that it takes to detect n photons is then described by the Erlang distribution with a shape parameter given by n and the rate S. The average sampling time $\langle T\rangle$ is then given by:

$$\langle T\rangle = n/S\,. \tag{15.17}$$

For typical experiments, S = 100 kHz and n = 5, yielding a sampling time of $\langle T\rangle$ = 50 µs.

For BVA, it is assumed that the molecule is either in state 1 or 2 (with apparent FRET efficiencies $E_{\mathrm{app}}^{(1)}$ or $E_{\mathrm{app}}^{(2)}$) at each sampling point of the FRET efficiency. Here, we consider a two-state system with rates k_{12} and k_{21}:

$$E_{\mathrm{app}}^{(1)} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} E_{\mathrm{app}}^{(2)} \tag{15.18}$$

If the dynamics are fast, there is a certain probability that the molecule interconverts during the sampling time T, which reduces the observed variance. The average time for the molecule to convert from any state is given by the kinetic relaxation time τ_r:

$$\tau_r = \frac{1}{k_{12}+k_{21}} \tag{15.19}$$

The number of interconversions k during the sampling time T is Poisson distributed with the average value $\mu = T/\tau_r$:

$$P(k) = \frac{(T/\tau_r)^k}{k!}\,e^{-T/\tau_r}. \tag{15.20}$$

The probability to convert at least once during a time interval T is then given by:

$$P_{\text{convert}} = P(k \geq 1) = 1 - P(k = 0) = 1 - e^{-\frac{T}{\tau_r}}. \quad (15.21)$$

For simplification, we assume that, when an interconversion occurred, the sampling window converges to the average apparent FRET efficiency $\langle E_{\text{app}} \rangle$. Then, any sampling of the apparent FRET efficiency in which an interconversion occurred will have zero contribution to the variance and the dynamic contribution to the variance is simply reduced by the probability that no interconversion event occurred:

$$\text{Var}'_{\text{dyn}} = (1 - P_{\text{convert}})\text{Var}_{\text{dyn}} = e^{-\frac{T}{\tau_r}}\text{Var}_{\text{dyn}}. \quad (15.22)$$

By mixing the shot-noise and conformational contributions to the variance of the FRET efficiency, the dynamic FRET-line for BVA in the presence of fast conformational dynamics becomes:

$$\begin{aligned}\text{Var}\left(E_{\text{app}}, P_{\text{convert}}\right) &= (1 - P_{\text{convert}})\left(f_1\left[\frac{E_{\text{app}}^{(1)}\left(1 - E_{\text{app}}^{(1)}\right)}{n} + {E_{\text{app}}^{(1)}}^2\right]\right. \\ &\left. + (1 - f_1)\left[\frac{E_{\text{app}}^{(2)}\left(1 - E_{\text{app}}^{(2)}\right)}{n} + {E_{\text{app}}^{(2)}}^2\right]\right) + P_{\text{convert}}\frac{\langle E_{\text{app}}\rangle\left(1 - \langle E_{\text{app}}\rangle\right)}{n} - \langle E_{\text{app}}\rangle^2\end{aligned}$$

where f_1 and $(1 - f_1)$ are the fraction of time the molecule spends in state 1 and 2, respectively.

III. Discussion of detection limits

Using the described formalism, we tested whether hypothetical conformational dynamics between the apo and holo states of the studied systems, MalE-1 to 5 and U2AF, could be detected in a given situation. The FRET efficiencies of the apo and holo states were obtained from AV simulations on the available crystal structures as described in the online methods (Supplementary Table 19).

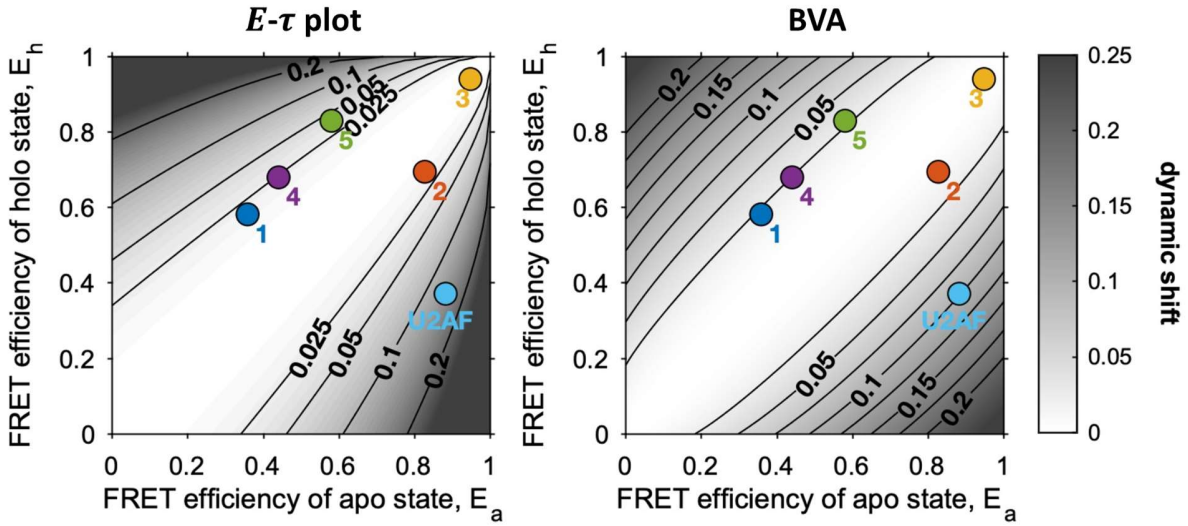
We first visualized the expected dynamic shift for the E-τ and BVA plots as a function of the FRET efficiencies of the limiting states (Supplementary Figure SN15.1). Along the diagonal, dynamic shifts are low (<0.025) because the FRET efficiency contrast is small whereas off-diagonal combinations of FRET efficiencies result in large dynamic shifts. While the BVA plot is symmetric between low and high FRET efficiencies, the E-τ plot is better at resolving a dynamic exchange at high FRET efficiencies of the limiting states compared to the low FRET efficiency region. This asymmetry arises because the uncertainty in the fluorescence lifetime is largest for long fluorescence lifetimes and thus for low FRET efficiencies.

The experimentally studied MalE mutants all showed dynamic shifts below 0.03 while, for U2AF, a dynamic shift of > 0.1 is expected for both the E-τ and BVA plots. To put these expected shifts

into perspective with respect to the experimental uncertainty, we plot the dynamic shift normalized to the expected standard error of the mean for different numbers of detected photons per burst and number of bursts in Supplementary Figure SN15.2 for the E-τ plot and Supplementary Figure SN14.3 for the BVA plot. Dynamics are detectable when the ratio of the dynamic shift to the experimental uncertainty, ds/σ_{SEM}, exceeds one. Clearly, the sensitivity to detect a dynamic exchange increases for both the E-τ and BVA plot with increasing number of photons and bursts. For typical experimental values (100 photons per burst and 1000 bursts), the dynamic exchange is predicted to be detectable as $ds/\sigma_{SEM} > 10$ for all experimental systems.

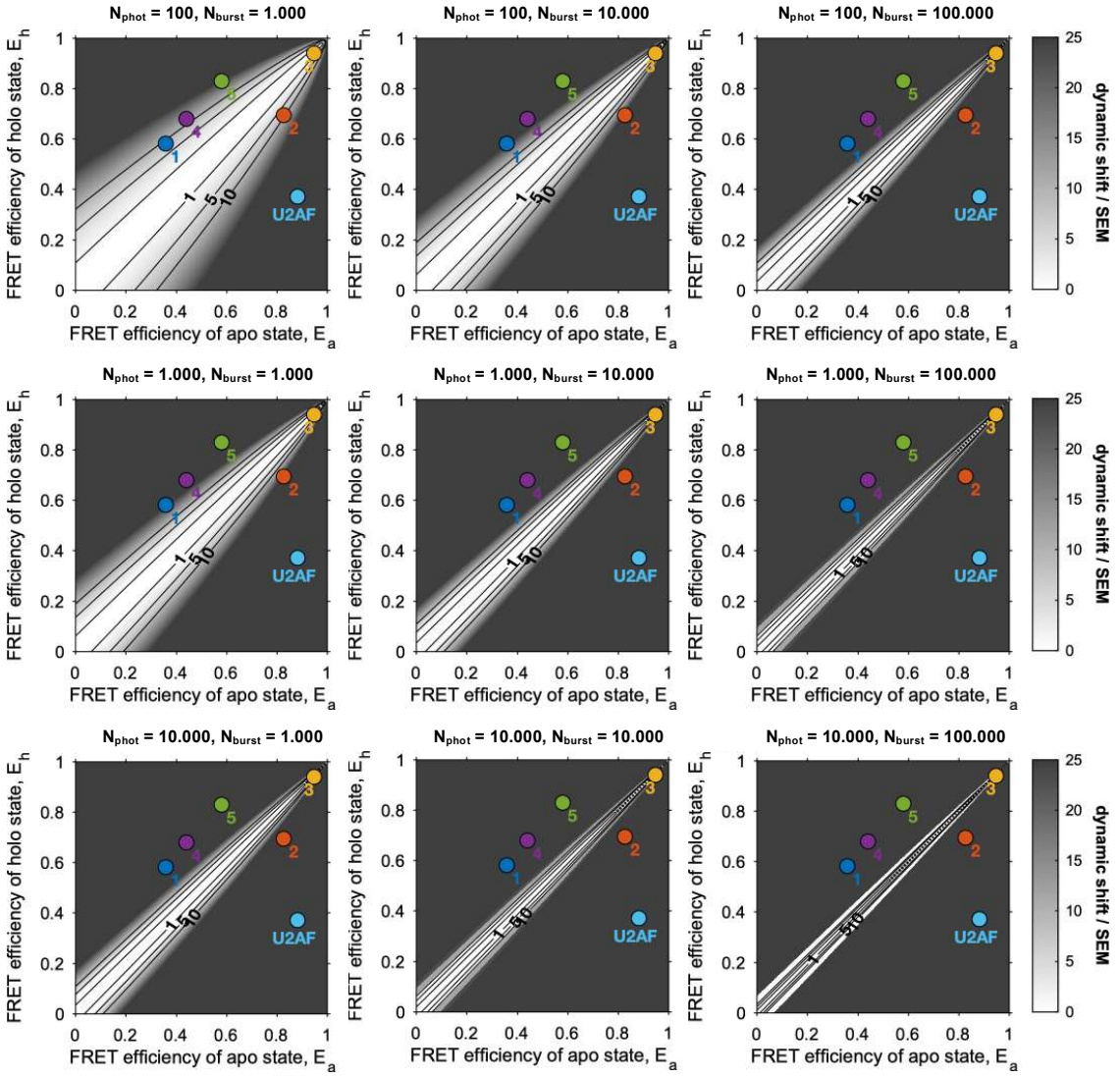
The sensitivity of the E-τ plot crucially depends on the accuracy of the correction factors used to compute accurate FRET efficiencies, in particular the detection efficiency correction factor γ. With higher uncertainty, $\Delta\gamma/\gamma$, the sensitivity of the E-τ plot decreases (Supplementary Figure SN15.4). At a relative uncertainty of $\Delta\gamma/\gamma = 0.1$, the sensitivity reduces to an extent that potential dynamics between the apo and holo states of the different MalE mutants would become undetectable. On the other hand, the large-scale dynamics for U2AF would remain detectable even at a high calibration uncertainty of $\Delta\gamma/\gamma = 0.3$. Note, the uncertainty discussed here is the uncertainty of the detection correction factor in a single lab and not the distribution of $\Delta\gamma/\gamma$ values calculated between labs in Fig. 3e.

The BVA plot loses sensitivity when the timescale of dynamics approaches the sampling window used for the estimation of the variance of the FRET efficiency distribution (Supplementary Figure SN15.5). For slow conformational dynamics with a relaxation time of τ_r = 10 ms, the potential exchange between the apo and holo states is detectable with $ds/\sigma_{SEM} > 10$. As the relaxation time approaches the sampling window (here, $T \approx 50$ µs), the sensitivity is reduced significantly. At τ_r = 20 µs, the exchange for most MalE mutants is on the border of the detection limit while, for τ_r = 10 µs, even potential dynamics of U2AF would become undetectable in BVA.

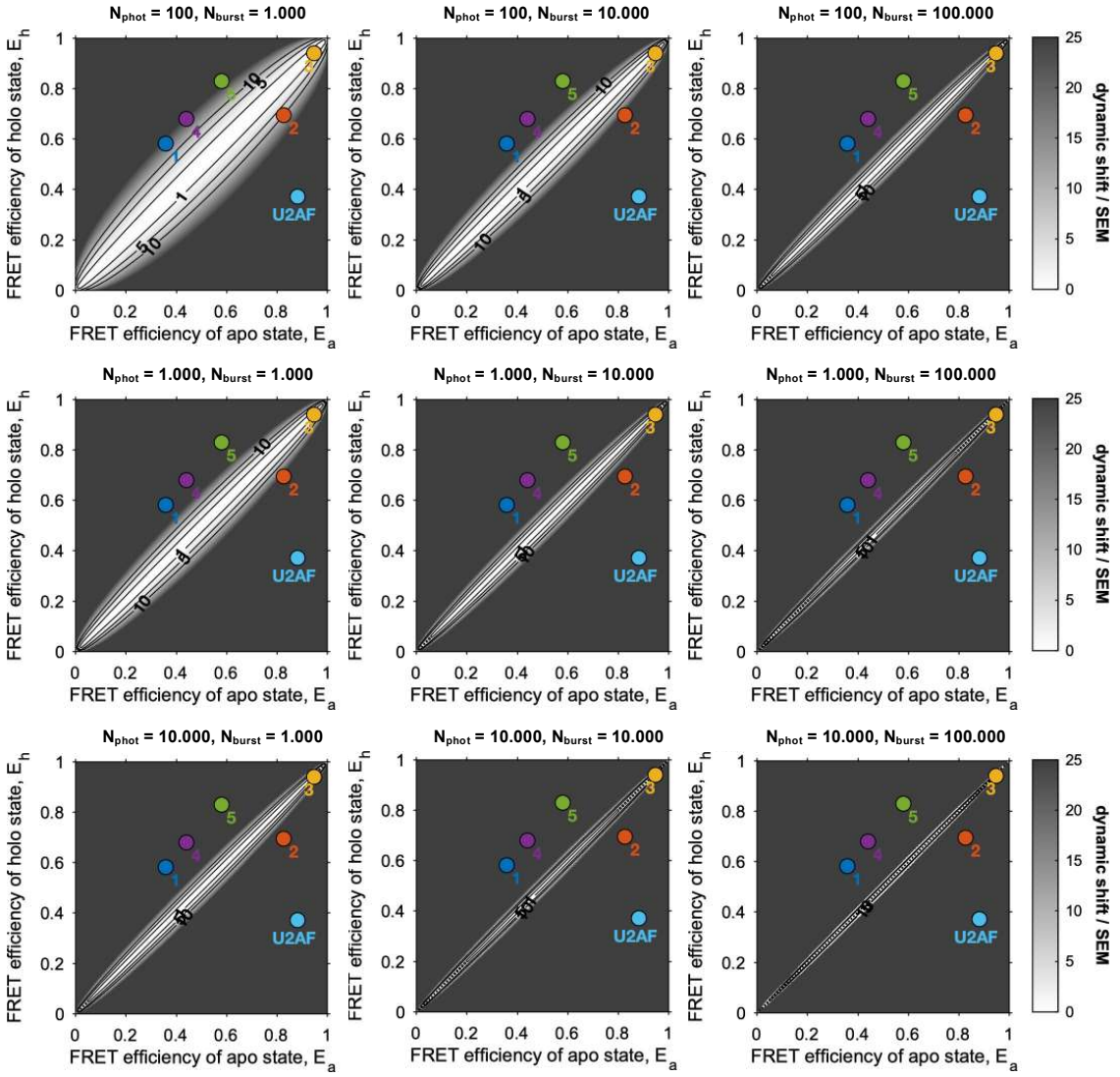


Supplementary Figure SN15.1: The dynamic shift in the E-τ (left) and the BVA (right) plots as a function of the FRET efficiencies of the limiting states (apo and holo) undergoing hypothetical dynamic exchange.

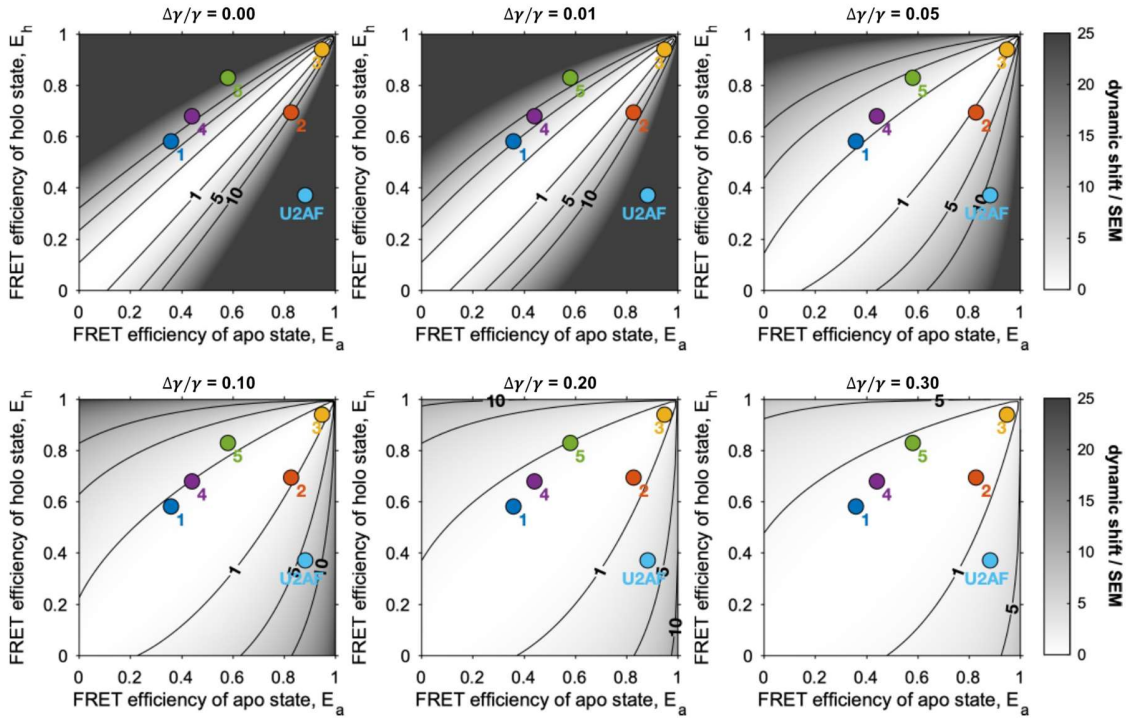
The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Table 19 and Supplementary Table 8).



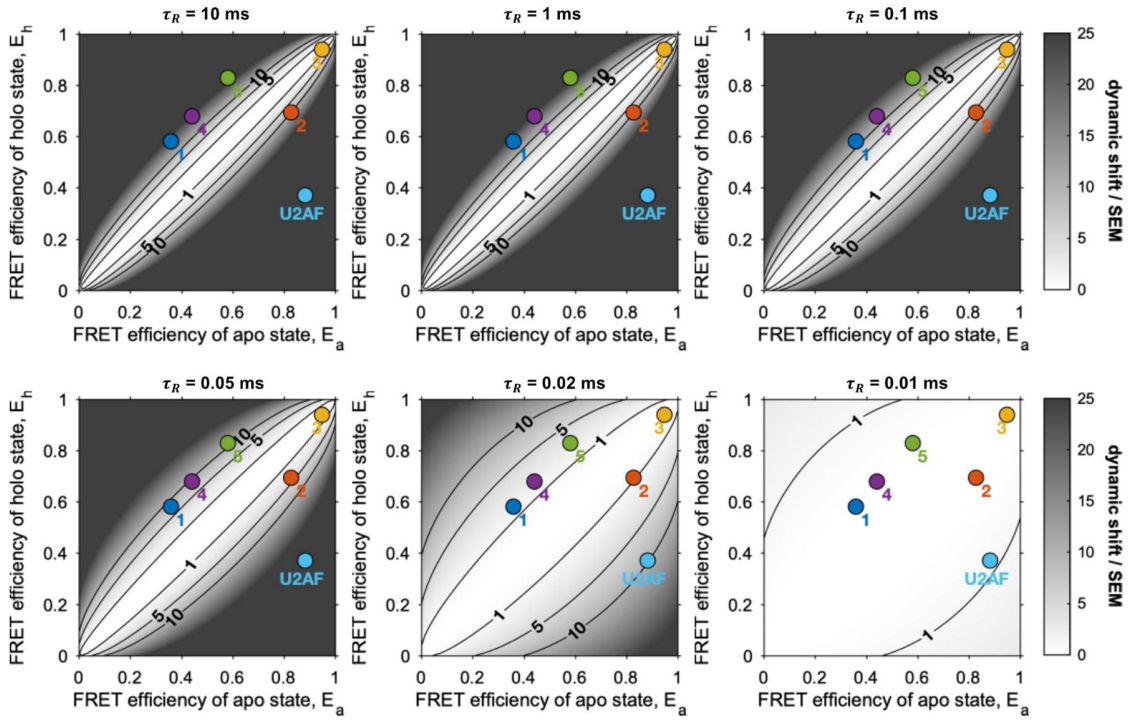
Supplementary Figure SN15.2: Detectability of conformational dynamics in the E-τ plot as a function of the photon-counting statistics. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Table 19 and Supplementary Table 8).



Supplementary Figure SN15.3: Detectability of conformational dynamics in the BVA plot as a function of the photon-counting statistics. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Table 19 and Supplementary Table 8).



Supplementary Figure SN15.4: Detectability of conformational dynamics in the E-τ plot as a function of the calibration error of the detection efficiency correction factor γ. The relative uncertainty of the γ-factor is given by $\Delta\gamma/\gamma$. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). It is assumed that the dynamic population contains 1000 bursts of 100 photons. The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Table 19 and Supplementary Table 8).



Supplementary Figure SN15.5: Detectability of conformational dynamics in the BVA plot as a function of the time scale of the dynamic exchange, quantified by the relaxation rate τ_r. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). It is assumed that the dynamic population contains 1000 bursts of 100 photons and a count rate of 100 kHz, and that an averaging window of 5 photons is used for the BVA analysis. The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Table 19 and Supplementary Table 8).

Supplementary Note 16: Model-free analysis of fluorescence decays for U2AF65.

Sub-ensemble fluorescence decays. Sub-ensemble donor fluorescence decays were generated from single molecule experiments by selecting the double-labeled population using the ALEX-2CDE filter[9] with an upper limit of 15. An additional stoichiometry cut was applied ($0.4 \leq S \leq 0.55$) to remove the dye-related artifact at high FRET efficiency that showed a higher stoichiometry of ~0.6 (Supplementary Figure 15). For the donor-only reference decays, a stoichiometry threshold of $S \geq 0.98$ was used.

Fitting procedure. The ideal fluorescence decays $f(t)$ were convoluted with the instrument response function and corrected for background contributions to the TCSPC pattern due to, for

example, the contribution of scattered laser light, autofluorescence of the buffer, dark counts and uncorrelated background signal, to obtain the model decay $F(t)$:

$$F(t) = F_0 \cdot f(t) \otimes IRF(t - t_{\text{shift}}) + sc\, BG(t), \tag{16.1}$$

where $\otimes$ here denotes a linear convolution, F_0 is the initial amplitude of the fluorescence decay, $IRF(t)$ is the instrument response function (which is shifted by the time t_{shift}), $BG(t)$ is the normalized background/scatter pattern obtained from a buffer measurement and sc is the background/scatter amplitude, which is estimated from the cumulative duration of all analyzed bursts T_{bursts} and the background count rate c_{BG} as $sc = c_{BG} T_{\text{bursts}}$. Note that, contrary to Supplementary Note 9, we apply the linear convolution operation here for the analysis of smFRET data recorded using PIE because the analysis is performed only a selected time interval of the microtime histogram (i.e., a PIE channel) to restrict the analysis to the fluorescence decay of the donor. Hence, the signal here is not periodic as it is in the ensemble TCSPC experiments.

Fits were optimized using the reduced chi-square defined by:

$$\chi^2_{\text{red}} = \frac{1}{N_{\text{data}} - N_{\text{param}}} \sum \frac{\left(F_{exp}^{(k)} - F_{model}^{(k)}\right)^2}{w_k^2}, \tag{16.2}$$

where $F_{exp}^{(k)}$ and $F_{model}^{(k)}$ are the intensities of the measured and model decay in the TCSPC bin k, w_k is the respective weight given by $w_k = \sqrt{F_{exp}^{(k)}}$ based on the Poisson statistics of the detected signal, and N_{data} and N_{param} are the number of bins in the TCSPC histogram and the number of independent fit parameters, respectively.

Pre-fitting to estimate the background parameters. In a first step, we performed a global fit of the donor-only fluorescence decay $F_{DO}(t)$ and the FRET-induced donor decay $F_{DA}(t)$ using a two-component Gaussian distribution for the interdye distance. This allows us to estimate the fluorescence lifetimes and respective amplitudes of the donor-only sample as well as the parameters t_{shift} and F_0 of the FRET-induced donor decay. The donor-only decay is described using two lifetime components:

$$f_{DO}(t) = \sum_{i=1}^{2} x_{D(0)}^{(i)} \exp\left(-t/\tau_{D(0)}^{(i)}\right), \tag{16.3}$$

where $\tau_{D(0)}^{(i)}$ and $x_{D(0)}^{(i)}$ are the lifetime and fraction of donor-only species i. The presence of the acceptor acts as an additional process that depopulates the donor excited state at a rate of:

$$k_{RET}(R_{DA}) = \frac{1}{\tau_{D(0)}} \left(\frac{R_0}{R_{DA}}\right)^6, \tag{16.4}$$

where R_0 is the Förster radius and R_{DA} is the donor-acceptor separation distance. Note that $\tau_{D(0)}$ refers here to the lifetime of the species with the respective quantum yield that is used for the

calculation of R_0, which can be different from the $\tau_{D(0)}^{(i)}$ obtained for the donor-only decay. For a given distance distribution, $p(R_{DA})$, the FRET-induced donor decay is then given by:

$$f_{DA}(t) = \left(1 - x_{\mathrm{DOnly}}\right) f_{DO}(t) \left(\int p(R_{DA}) \exp[-k_{RET}(R_{DA})\, t]\, dR_{DA} \right) + x_{\mathrm{DOnly}} f_{DO}(t), \quad (16.5)$$

where x_{DOnly} is the contribution of a donor-only signal due to acceptor photoblinking or photobleaching.

The two-component Gaussian distance distribution is given by:

$$p_{2G}(R_{DA}) = \sum_{i=1}^{2} x_{DA}^{(i)} \left(\sqrt{2\pi}\sigma_{DA,i}\right)^{-1} \exp\left[-\frac{\left(R_{DA} - \langle R_{DA}^{(i)} \rangle\right)^2}{2\sigma_{DA,i}^2} \right], \quad (16.6)$$

where $x_{DA}^{(i)}$ is the amplitude, $\langle R_{DA}^{(i)} \rangle$ the average interdye distance and $\sigma_{DA,i}$ the width of component i. The model is globally optimized with respect to the amplitude and lifetimes of the donor-only components. All parameters, except for the distance distribution $p(R_{DA})$, were fixed for the maximum entropy method model-free analysis discussed below.

Model-free analysis. The maximum entropy method (MEM) is an approach to extract the most unbiased distribution of a given parameter that provides a satisfactory fit to the experimental data[38–40]. Instead of minimizing the reduced chi-square, χ_{red}^2, the following functional is maximized:

$$\Theta = vS - \chi_{\mathrm{red}}^2, \quad (16.7)$$

where v is a constant scaling factor and S is the entropy functional of the parameter distribution. The entropy, S, of a discrete probability distribution p_i is defined by:

$$S = -\sum_i p_i \log \frac{p_{\mathrm{i}}}{m_i}, \quad (16.8)$$

where p_i is the distribution of the parameter of interest and m_i describes the prior knowledge of the parameter distribution. We applied the MEM analysis to extract the distribution of interdye distances R_{DA}, $p(R_{DA})$. The measured FRET-induced donor fluorescence decay $f_{DA}^{exp}(t)$ is described as a superposition of exponential functions given by:

$$f_{DA}^{exp}(\mathrm{t}) = \left(1 - x_{\mathrm{DOnly}}\right) f_{DO}(t) \left(\sum_j p\left(R_{DA}^{(j)}\right) \exp\left[-k_{RET}\left(R_{DA}^{(j)}\right) t \right] \right) + x_{\mathrm{DOnly}} f_{DO}(t) \quad (16.9)$$

where the summation is performed over a range of interdye distances $R_{DA}^{(j)}$ from 10 to 150 nm. Maximization of Θ is performed as described in Vinogradov and Wilson[41] over a wide range of values for the regularization parameter v. The choice of the regularization parameter v was done by visual inspection of the L-curve plot of the negative entropy, $-S$, against the reduced chi-squared χ_{red}^2. The resulting values for χ_{red}^2 and the regularization parameter v are given in Supplementary Table SN16.1. All analyses were performed using the *TauFit* module of the PAM software package[8].

Prior distribution. The prior distribution is based on the detached apo ensemble derived in Huang et al.[42] For each of the 200 structures in the ensemble, the average interdye distance $\langle R_{DA} \rangle$ was determined using AV simulations[23,43]. To account for additional broadening due to the flexible dye linker, a kernel density estimate using a Gaussian kernel with a fixed width of 6 Å was performed to obtain the prior distribution as shown in Fig. 6d of the main text. In addition, we performed a kernel density estimation of the interdye distance distribution using a Gaussian kernel by the *ksdensity* function of MATLAB, which computes the theoretically optimal bandwidth for normally distributed data[44]. This procedure returned a similar bandwidth of 5.7 Å. The different histograms and the kernel density estimate of the detached apo ensemble are compared in Supplementary Figure SN16.1.

Deconvolution of the probability distribution obtained by the MEM analysis. The obtained distribution of the donor-acceptor distance R_{DA} from the MEM analysis is broadened due to the flexible dye linkers. The magnitude of this additional broadening has previously been characterized to be on the order of ~6 Å[23]. To obtain the distribution of mean donor-acceptor distances $\langle R_{DA} \rangle$, we hence perform a deconvolution of the R_{DA} distribution obtained by MEM.

The deconvolution is performed using a Gaussian kernel with a width of σ_{DA} = 6 Å, defined on a distance grid of $\langle R_{DA} \rangle \in [10,150]$ with a resolution of 0.5 Å. The kernel matrix $\boldsymbol{Q}$ is defined as:

$$\boldsymbol{Q} = \begin{pmatrix} g_Q\left(R_{DA}^{(1)}; \langle R_{DA} \rangle^{(1)}\right) & \cdots & g_Q\left(R_{DA}^{(1)}; \langle R_{DA} \rangle^{(M)}\right) \\ \vdots & \ddots & \vdots \\ g_Q\left(R_{DA}^{(N)}; \langle R_{DA} \rangle^{(1)}\right) & \cdots & g_Q\left(R_{DA}^{(N)}; \langle R_{DA} \rangle^{(M)}\right) \end{pmatrix}, \tag{16.10}$$

where the indices N and M represent the number of sampling points of the discrete distributions for R_{DA} and $\langle R_{DA} \rangle$, respectively, and the kernel functions are given by:

$$g_Q\left(R_{DA}^{(n)}; \langle R_{DA} \rangle^{(m)}\right) = \left(\sqrt{2\pi}\sigma_{DA}\right)^{-1} \exp\left[-\frac{\left(R_{DA}^{(n)} - \langle R_{DA} \rangle^{(m)}\right)^2}{2\sigma_{DA}^2}\right].$$

Given the discretized probability distribution of mean donor-acceptor distances $\langle R_{DA} \rangle$, expressed as the row vector:

$$\boldsymbol{p}_{\langle R_{DA} \rangle} = \begin{pmatrix} p_{\langle R_{DA} \rangle}^{(1)} \\ p_{\langle R_{DA} \rangle}^{(2)} \\ \vdots \end{pmatrix}, \tag{16.11}$$

the discrete distribution of R_{DA}, $\boldsymbol{p}_{\boldsymbol{R_{DA}}}$, is obtained as:

$$\boldsymbol{p}_{\boldsymbol{R_{DA}}} = \boldsymbol{Q} \cdot \boldsymbol{p}_{\langle \boldsymbol{R_{DA}} \rangle}. \tag{16.12}$$

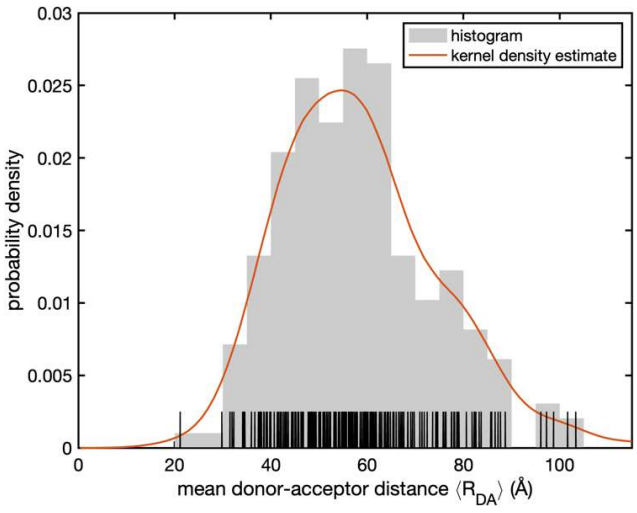
To obtain an estimate of $\boldsymbol{p}_{\langle \boldsymbol{R_{DA}} \rangle}$, we minimized the absolute value of the difference between the measured distribution $\boldsymbol{p}_{\boldsymbol{R_{DA}}}^{(\mathbf{exp})}$ and the distribution obtained by Eq. 16.12 above:

$$\min \left\| \boldsymbol{p}_{\boldsymbol{R_{DA}}}^{(\mathbf{exp})} - \boldsymbol{Q} \cdot \boldsymbol{p}_{\langle \boldsymbol{R_{DA}} \rangle} \right\|, \tag{16.13}$$

under the constraint that all elements of $\boldsymbol{p}_{\langle R_{DA}\rangle}$ must be positive. In addition, we used the kernel density estimate of the distribution of $\langle R_{DA}\rangle$ as obtained from the detached apo ensemble as a starting point for $\boldsymbol{p}_{\langle R_{DA}\rangle}$. The optimization is performed using the *fmincon* function of MATLAB. The deconvolved $\boldsymbol{p}_{\langle R_{DA}\rangle}$ distributions are shown in Fig. 6d of the main text.

Supplementary Table SN16.1: Reduced chi-squared χ^2_{red} and regularization parameter v for the model-free analysis of fluorescence decays of U2AF65 measured in Lab #2.

	Dye Pair					
	Atto532-Atto643		**Alexa546-Alexa647**		**Alexa488-Alexa647**	
Sample	$\chi^2_{\mathrm{red.}}$	v	$\chi^2_{\mathrm{red.}}$	v	$\chi^2_{\mathrm{red.}}$	v
U2AF65 apo	1.30	5.09	1.28	3.20	1.29	2.21
U2AF65 holo	1.03	1.05	1.24	1.05	1.13	1.05



Supplementary Figure SN16.1: Probability distribution of the mean donor-acceptor separation obtained for the detached apo-ensemble of U2AF65 as reported in Huang et al.[42] For each of the 200 structures in the ensemble, the mean donor-acceptor distance $\langle R_{DA}\rangle$ for the dye pair Atto532-Atto643 was computed from AV simulations. The vertical black lines indicate the individual values of $\langle R_{DA}\rangle$ and the gray bars shows the distance histogram computed with a bin width of 5 Å. The kernel density estimate (red line) was computed using a Gaussian kernel with a bandwidth of 6 Å, providing a smooth estimate of the probability density that is in good agreement with the histogram. The kernel density estimate was used as the prior distribution for the maximum entropy analysis of the fluorescence decays (Fig. 6d). See Supplementary Note 16 for details.

Supplementary Note 17: Filtered-FCS Analysis of the U2AF65 kinetics.

Filtered-FCS (fFCS) calculates the correlation functions for distinct species using statistical weighting based on the TCSPC patterns for the different species[45]. For the fFCS of the apo state of U2AF65 shown in Fig. 6e in the main text, the correlation functions were calculated as described in Barth et al.[46] using the PAM software package[8]. Briefly, two species with low and high FRET efficiency (LF, HF) were selected based on the FRET efficiency thresholds of $E \leq 0.6$ and $E \geq 0.9$, respectively. Filter functions for the fFCS analysis were calculated based on the concatenated TCSPC patterns of the donor and FRET-sensitized acceptor fluorescence decays. As a third species, a background pattern from a buffer measurement was included to account for the contribution of scattered laser light and constant background signal. "Purified" fFCS correlation functions were computed only for the detected and filtered double-labeled bursts. Signal occurring within 5 ms before or after the burst was included to obtain a better estimate of the diffusional part of the correlation function. To avoid detector afterpulsing in the species autocorrelation functions, the signal detected in the parallel and perpendicular channels of the MFD setup was filtered separately and cross-correlated. The two species autocorrelation functions (SACF) and two species cross-correlation functions (SCCF) were globally analyzed using a model with one diffusion term and two kinetic terms:

$$\begin{aligned} G(t_c) &= G_{\text{diff}}(t_c) G_{\text{kin}}(t_c) \\ G_{\text{diff}}(t_c) &= \frac{1}{\langle N \rangle}\left(1 + \frac{t_c}{t_{\text{diff}}}\right)^{-1}\left(1 + \frac{t_c}{p\, t_{\text{diff}}}\right)^{-\frac{1}{2}} \\ G_{\text{kin}}^{SACF}(t_c) &= 1 + A_1 e^{-\frac{t_c}{t_{R,1}}} + A_2 e^{-\frac{t_c}{t_{R,2}}} \\ G_{\text{kin}}^{SCCF}(t_c) &= 1 - A_1 e^{-\frac{t_c}{t_{R,1}}} - A_2 e^{-\frac{t_c}{t_{R,2}}} \end{aligned} \tag{17.1}$$

Here, $G_{\text{diff}}(t_c)$ and $G_{\text{kin}}(t_c)$ are the diffusion and kinetic part of the correlation function $G(t_c)$ at lag time t_c. For the diffusion component, $\langle N \rangle$ is the average number of particles in the observation volume, t_{diff} is the diffusion time and the structural factor p is given by the ratio of the axial to the lateral width of the observation volume (typically, p = 5-10). For the kinetic component, $t_{R,1}$ and $t_{R,2}$ are the kinetic relaxation times with amplitudes A_1 and A_2 respectively. For the analysis, the kinetic relaxation times and diffusion time are optimized globally over all four curves. The amplitude of the diffusional part (given by $\langle N \rangle^{-1}$) and the kinetic terms (A_1 and A_2) were optimized individually for each of the four curves. This is necessary as the amplitude information in fFCS is less reliable compared to the time evolution of the curves due to the effect of imperfect filters[45].

Supplementary Note 18: Dynamics of U2AF65.

From previous work, we were aware that U2AF65 is a dynamic protein system whose conformational distribution can be shifted by the addition of an RNA ligand[47]. From the BVA and *E*-τ plots (Fig.4), it is evident that the apo protein is dynamic on the sub-millisecond timescale. The *E*-τ plot already contains information regarding the timescale of the dynamics of the sample with respect to the duration of the burst. A single transition is sufficient to cause a shift for that burst from the static FRET line in the *E*-τ plot. Hence, dynamics that are up to an order of magnitude slower than the diffusion time can still be detected. In this case, individual "static" populations should be observable on the static FRET line with a smearing between the states coming from the few bursts where transitions occurred during the burst. For faster dynamics, several transitions occur during a burst and the individual "static" populations disappear with only a single state observable on the dynamic FRET line. The position of the population along the dynamic FRET line depends on the equilibrium between the different states[14]. For U2AF65, the dynamics are fast enough that we only observed an average FRET state and the individual populations are not visible.

To get an idea of the distribution of FRET efficiencies in the sample, as is known from NMR and SAXS data, there is not a single closed or a single open conformation, but a family of conformations[42]. To this end, we performed a model-free analysis of the donor fluorescence lifetime by the maximum entropy method[38–40] to infer the underlying distribution of interdye distances (Fig. 6d in the main text, Supplementary Note 16).

After establishing the distribution of FRET efficiencies within the sample, we next determined the timescale of the dynamics. To get a model-independent estimation of the timescale of kinetics faster than the burst duration, we utilized an FCS approach. One possibility to visualize the dynamics is by performing a cross-correlation analysis between the donor and acceptor signals (FRET-FCS)[48,49]. We choose to use filtered-FCS (fFCS)[45,50] (Supplementary Note 17) as it increases the contrast of the FCS signal in comparison to FRET-FCS by weighing the correlation signals depending on the difference in the lifetime, color and anisotropy of fluorescence fluctuations between the bursts showing the lowest and highest FRET values. For the Atto532-Atto643 labeled protein, we observed two relaxation times with values of ~ 10 µs and ~ 200 µs (Fig. 6e, Supplementary Table 14). We assign the faster kinetics to dynamics of the detached domains and the slower rate to interconversion between compacted conformations within the conformational ensemble.

A second approach for extracting the dynamic timescales from a sample is to use the dynamic photon distribution analysis (PDA) method[13]. Dynamic PDA allows one to delineate the conformational heterogeneity over sub-millisecond transitions from the width of the FRET histograms beyond shot-noise by analyzing the raw photon counts[51,5213]. The FRET efficiency histograms are calculated using different time-windows for the binning and then fit using a global model (Fig. 6f in the main text and Supplementary Figure 14). There is little change in the FRET efficiency histograms when varying the integration time between 0.5 ms, 1.0 ms, 1.5 ms and 2.0

ms. From a global fit to the different histograms for the apo and holo conditions, the states as well as the conversion rates between them (provided they are between ca 100 µs and the timescale of diffusion) can be extracted. To simplify the complexity of the dynamic PDA model, we used a broad Gaussian function to empirically describe the fast, dynamic ensemble observed with fFCS and used dynamic PDA to look for slower kinetics. We observed a slow interconversion on the timescale of 10-100 ms between the very-high-FRET state at $E \sim 0.95$ and the main population at $E \sim 0.85$. The high FRET population (E = 0.85) is clearly distinct from a second peak in the E-τ plot of E ~ 0.95 and has a very short donor lifetime. The very-high-FRET signal is reminiscent of what is observed for dye-dye interactions[53]. If this is true, the rates given by dynamic PDA should change when a different dye-pair is used whereas the protein dynamics should remain unchanged. Indeed, when evaluating measurements using the Alexa546 – Alexa647 dye pair, the fast dynamics remains unchanged (90 – 200 µs) whereas the amplitude of this very-high-FRET peak and the rate of the slow dynamics changes slightly (Supplementary Figure 17 and Supplementary Table 17). This demonstrates the importance of performing smFRET measurements with at least two different dye-pairs[53]. However, for this analysis, the high stoichiometry state at high FRET efficiency was excluded suggesting the measured dynamics investigate real conformational changes of the protein.

Due to the complexity of the dynamics, we did not ask the various groups to provide a detailed analysis of the kinetics. We did ask laboratories, when possible, to provide an estimate of the timescale of the dynamics, when present. Five labs could contribute to a dynamic quantification of U2AF65 (Supplementary Table 14). These all correctly reported a quasi two-state behavior of the U2AF65 system. A comparison of kinetic rates based on further evaluation with dynamic PDA and fFCS for U2AF65 from different labs showed good consistency. Specifically, the calculated relaxation times for the apo state evaluated with fFCS were consistent with an approximately two-fold variation across labs (~200, ~320 and ~370 µs from three different labs). A dynamic PDA analysis for the holo-state kinetic rate estimations were not fully consistent most likely due to the variation seen in the FRET histograms (perhaps due to differences in temperature). Three among the five groups provided the kinetic rates for the holo-state with only ~20% variation of the reported relaxation times ($\tau_R = 1/(k_{12}+k_{21})$) of 1.25 ms, 1.42 and 1.6 ms (see Supplementary Table 14).

Challenges with the U2AF65 sample and analysis:

From the detailed analysis described above, it is clear that U2AF65 is not a simple two-state dynamic system. Apo U2AF65 exhibits dynamics on the microsecond timescale between a family of open and of closed conformations. Both apo and holo measurements exhibit a slow dynamic exchange between a very-high-FRET populations (at $E \sim 0.95$) and the other FRET states.

Beyond the complicated dynamics, other factors also impacted the analysis of the dynamic state. First of all, the measured RNA ligand concentration was not high enough to saturate binding to U2AF65. From the measured binding affinity of $K_d \approx 1.2$ µM (Supplementary Figure 6), the ligand concentration of 5 µM used in these measurements lead to approximately 85% of the sample having a ligand and 15% remaining in the apo conformation. This corresponds well to the measured ratio of the two populations. Alternatively, the protein could theoretically exchange between a quasi-

static state and a dynamic conformation where one of the RRM domains transiently releases the RNA strand. However, there is no evidence that the RNA-bound state of U2AF65 interconverts with the apo-like conformation on the timescale of a burst (~ 10 ms). This is consistent with what is expected assuming a K_d of 1.2 µM and a measured k_{on} of 0.7 μM^{-1} s^{-1}, leading to a k_{off} of ~ 1s. This timescale is too slow to be detected with the solution-based measurements. Hence, we attributed the dynamics detected by BVA and the E-τ plots in the holo measurements to the significant presence of proteins not having an RNA ligand bound (Fig.5). Also, the timescale of the kinetics for the dynamic population was similar to that of the apo-state (~ 100 - 200 µs). The second FRET peak with a FRET efficiency of E ~ 0.44 observed in the presence of RNA is close to the static line suggesting that the RNA bound conformation is static.

A second challenge in analyzing the U2AF65 data comes from the fraction of molecules in the very-high-FRET state. Due to dye-dye interactions, the fluorescence intensity yields a different stoichiometry value and influences determination of the γ detection correction factor when it is not excluded from the analysis (Supplementary Figure 15).

In summary, apo U2AF65 appears to undergo fluctuations between families of open and closed states with dynamics on the timescale of 100-200 µs. Dynamics on the millisecond timescale are also observed due to the formation of stable compacted states. Fluctuations between an open and a closed conformation is a fair approximation of the dynamics, but an ensemble of open and closed conformations are needed to fully describe the measured data as shown in Fig. 6c-f.

Supplementary Note 19: Overview of set-ups and analysis software used across all labs

Lab#1

All sample solutions were measured in Lab-Tek I chamber slides at a concentration of ~50-100 pM. Single-molecule FRET experiments with MFD-PIE were performed on a homebuilt confocal setup as described previously[4].

For samples labeled with Alexa546-Alexa647 and Atto532-Atto643, the our two-color green-red setup was used: Fluorescent donor molecules were excited by an amplified, frequency-doubled diode laser (PicoTA 530, PicoQuant, Berlin) at 532 nm and acceptor molecules were excited by a pulsed diode laser (LDH-D-C 640, PicoQuant) at 640 nm. Both lasers were operated at a power of 50 µW measured at the sample, operated at a repetition rate of 26.66 MHz and synchronized with a delay of 18 ns. The laser light was guided into the epi-illuminated confocal microscope base (Nikon Eclipse TE300) and focused by a 60X water immersion objective (Plan Apo IR 60x/1.27, Nikon). For excitation, only parallel and perpendicular polarized light was selected using a Glan-Thompson polarizer (GTHM Polarizer, Thorlabs) before focusing into the objective. The emitted fluorescence was collected through the objective and spatially filtered using a pinhole with 75 µm diameter. and spectrally split . The fluorescence emission was first separated for polarization using a polarizing beam-splitter (PBS3, Thorlabs) before being spectrally separated into donor and acceptor channels by a dichroic mirror (640DCXR; AHF Analysentechnik). Fluorescence emission was filtered (donor: Brightline HQ582/75, acceptor: Brightline HQ700/75, AHF Analysentechnik) and focused on avalanche photodiodes (SPCM-AQR, Perkin-Elmer) for detewction. The detector outputs were recorded using four TCSPC cards (SPC154; Becker and Hickl)

For samples labeled with Alexa488-Alexa647, our three-color setup was used as described previously[54]. For this dye-pair combination, we switched off the 560 nm laser and associated detectors. In general, the three-color set-up has three pulsed lasers with a ~20 ns delay between the laser pulses (pulse frequency of 16.7 MHz) (PicoQuant, Germany; LDH-D- C-485, LDH-D-TA-560, LDH-D-C-640). The lasers were synchronized using a laser driver (PicoQuant, Germany; Sepia II). A 60x water immersion objective with 1.27 N.A. (Nikon, Germany; Plan Apo IR 60x 1.27 WI) was used for focusing the lasers into the sample. The measured laser powers before the objective were ~120 µW for blue, ~75 µW for green and ~35 µW for red laser. Fluorescence was collected by the same objective, separated from the laser excitations using a polychroic mirror (AHF Analysentechnik; zt405/488/633, Germany) and confocal geometry was achieved using a 50 µm pinhole. Light passing through the pinhole was further separated by a polarizing beam splitter (Thorlabs, Germany). Separation of blue and red wavelengths were performed with a dichroic mirror (AHF Analysentechnik; 640DCXR). Emission filters (AHF Analysentechnik; ET525/50, ET670/30) were placed right before the APD detectors (LaserComponents, 2x COUNT-100B for blue detection; Perkin Elmer, 2x SPCM-AQR14 for red detection). Photons were recorded and synchronized to the lasers pulses using a TCSPC module (PicoQuant; HydraHarp400).

Data analysis was performed using the PAM (PIE Analysis with Matlab) software package as described elsewhere[8]. Single-molecule events were identified using a sliding time window analysis with a threshold of 50 photons, a time window of 500 µs and a minimum photon number of 10. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter was applied using an upper threshold of 12.

Lab#2

All sample solutions were measured in NUNC chambers (Lab-Tek, Thermo Scientific) with 500 µL sample volume and a pM concentration. Single-molecule FRET experiments with PIE were performed on a homebuilt confocal setup as described previously[22].

Setup #1: for samples labelled with Alexa Fluor 546 – Alexa Fluor 647 and Atto 532-Atto 643: The fluorescent donor molecules are excited by a pulsed white light laser source (SuperK EXTREME, NKT Photonics), using a modulator (SuperK Varia, NKT Photonics), operated at 25 MHz, 80 µW. The acceptor molecules are excited by a pulsed diode laser (LDH-D-C 640), operated at 25 MHz and 10 µW. Laser powers were measured at objective. Laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dichroic beamsplitter F68-532_zt532/640NIRrpo (AHF, Germany) and focused on a sample by a water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 100 µm diameter and further split into parallel and perpendicular components using polarizing beam splitter cube (VISHT11, Gsänger). Light is then spectrally split into "green" and "red" spectral windows by a dichroic mirror (T640lpxr, AHF, Germany). Fluorescence emission was filtered (donor: 47-595/50 ET, acceptor: HQ 730/140, AHF, Germany) prior to detection using avalanche photodiodes (SPCM-AQRH 14, Excelitas). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant).

Setup #2: for samples labelled with Alexa488-Alexa647: Donor molecules were excited by a pulsed diode laser (LDH-D-C 485, PicoQuant) at 485 nm. Acceptor molecules were excited with pulsed diode laser (LDH-D-C 640, PicoQuant) at 635 nm. Lasers were operated with the repetition frequency of 32 MHz, and with a delay with respect to each other of 10.5 ns. Laser powers were measured at the objective and were 60 µW for donor excitation laser and 10 µW for acceptor excitation laser. Laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dichroic beamsplitter FF500/646-Di01 (Semrock, USA), and focused on the sample by a water immersion objective (UPlanSApo 60x/1.2 NA, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and focused on a 100 µm pinhole. Using a polarizing beam splitter cube, emitted light is divided into its parallel and perpendicular components. This is then followed by light being split into two spectral windows, "green" and "red", using longpass filter Q595, and then again using 50/50 beam splitters resulting in a total of eight detection channels. Additionally, bandpass filters are placed in front of the detectors (FF01-530/43-25; AHF, Tübingen, Germany for donor molecules and HQ 720/150 nm; AHF, Tübingen, Germany for acceptor molecules). Detection is performed using eight avalanche photodiodes (4 green channels: τ-SPAD (PicoQuant, Germany) and 4 red channels: SPCM-AQR-14 (Perkin Elmer). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant). For both setups data analysis was performed using home-written LabView-based software. Burst search was performed as described[55] ,using APBS (All Photon Burst Search) method and inter-photon times as threshold.

Lab#3

Sample solutions were measured with 100 µl drop on coverslip with concentration of around 50 pM. Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal microscope as described previously[56]. The fluorescent donor molecules are excited by a diode laser OBIS 532-100-LS (Coherent, USA) at 532 nm operated at 60 µW at the sample in alternation mode. The fluorescent acceptor molecules are excited by a diode laser OBIS 640-100-LX (Coherent, USA) at 640 nm operated at 25 µW at the sample in alternation mode (100 µs alternation period). The lasers are combined by an aspheric fiber port (PAF2S-11A) and coupled into a polarization maintaining single-mode fiber P3-488PM- FC-2 (Thorlabs, USA). The

laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dual-edge beamsplitter ZT532/640rpc (Chroma/AHF) focused by a water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror H643 LPXR (AHF). Fluorescence emission was filtered (donor: BrightLine HC 582/75 (Semrock/AHF), acceptor: Longpass 647 LP Edge Basic (Semroch/AHF), focused on avalanche photodiodes (SPCM-AQRH-64, Excelitas). The detector outputs were recorded by a NI-Card PCI-6602 (National Instruments, USA).

Data analysis was performed using home written software package as described[56] .Single-molecule events were identified using a All-Photon-Burst-Search algorithm with a threshold of 15, a time window of 500 µs and a minimum total photon number of 150.

Lab#4

All sample solutions are measured in home build 60 µL chambers with ~ 100 pM concentration. PIE-FRET experiments are carried out on a home build confocal microscope. The fluorescent donor molecules are excited by a pulsed diode laser (LDH-P- FA-530B, PicoQuant), at 532 nm operated at 20 MHz, 55 µW at the sample in PIE experiment. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-D-C-640, PicoQuant), at 639 nm operated at 20 MHz, 50 µW at the sample in PIE experiment. The laserpulses are altered on the nanosecond timescale by a multichannel picosecond diode laser driver (PDL 828 “Sepia II”, PicoQuant GmbH) with an oscillator module (SOM 828, PicoQuant GmbH). The lasers were coupled into a single mode fiber (P3-488PM-FC, Thorlabs GmbH) to obtain a Gaussian beam profile. Circular polarized light was obtained by a linear polarizer (LPVISE100-A, Thorlabs GmbH) and a quarter-wave plate (AQWP05M- 600, Thorlabs GmbH). The laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dual-edge beam splitter (z532/633, AHF analysentechnik AG) focussed by an oil immersion objective (UPLSAPO100XO, NA 1.40, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (640DCXR, AHF Analysentechnik AG, Germany). Fluorescence emission was filtered (donor: Brightline HC582/75 (AHF Analysentechnik AG), RazorEdge LP 532 (Laser 2000 GmbH), acceptor: (Shortpass 750, AHF Analysentechnik AG; RazorEdge LP 647, Laser 2000 GmbH), focused on avalanche photodiodes (SPCM-AQRH-14-TR, Excelitas Technoligies GmbH & Co. KG). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant). The setup was controlled by a commercial software package (SymPhoTime64, Picoquant GmbH).

Data analysis was performed using PAM software package as described[8] . Single-molecule events were identified using a two channel APBS-algorithm with a threshold of 5 photons per time window, a time window of 500 µs and a minimum photon number of 20. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter was applied using an upper threshold of 10^9.

Lab#5

All sample solutions were measured in 200 µl PBS buffer with a labelled protein concentration of 25-100 pM. In short: Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal microscope as described previously[7,56,57] .The fluorescent donor molecules are excited by a spectrally

filtered laser beam of a pulsed supercontinuum source (SuperK Extreme, NKT Photonics) with an acousto-optical tunable filter (AOTFnc-VIS, EQ Photonics), at 532 nm and 640 nm. The laser light is guided into a single-mode fiber (PM-S405-XP, Thorlabs) and the collimated beam (Focusing collimator MB06, Q-Optics/Linos) was coupled into an oil-immersion objective (60×, NA 1.35, UPLSAPO 60XO, Olympus) by using a dichroic beam splitter (zt532/642rpc, AHF Analysentechnik) mounted on an inverse microscope body (IX71, Olympus). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (640DCXR, AHF Analysentechnik). Fluorescence emission was filtered (donor: Brightline HC582/75, acceptor ET700/75; AHF Analysentechnik), focused on avalanche photodiodes (Tau-SPAD, PicoQuant). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant).

Data analysis was performed as described[7,56,57]. Single-molecule events were identified using an All Photon Burst Search algorithm with a threshold of 15, a time window of 500 µs and a minimum photon number of 200.

Lab#6

Single molecule measurements were carried out on a home-build confocal microscope[58] .Pulsed green and red laser light (532nm, LDH-P-FA-530 and 640nm, LDH-D-C-640, respectively, PicoQuant) was polarised, overlaid and focused on the sample by an 60x water immersion objective (CFI Plan Apo VC 60XC/1.2 WI, Nikon). Excitation light was separated from the emitted light by a dichroic mirror (F53-534 Dual Line beamsplitter z 532/633, AHF). The emitted light was then guided through a further dichroic mirror (F33-647 beam splitter 640 DCXR, AHF) to separate donor and acceptor fluorescence. After spectral separation pinholes with a diameter of 150 mircometer refined the detection volume to about 8fL. Finally, the two photon streams were separated by polarizing beam splitters into their parallel and perpendicular parts and recorded by single-photon detectors (two SPCMAQR-14, PerkinElmer and two PDMseries APDs, Micro Photon Devices). Time-correlated single photon counting with picosecond resolution and data collection was performed by a HydraHarp400 (PicoQuant) and the Symphotime 32 software (PicoQuant).

Lab#7

All sample solutions were measured a drop on a coverslip, sealed in an airtight chamber, with concentration ~50 pM. The general scheme of the setup is described[59] . In short: Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal. The fluorescent donor molecules are excited by a continuous wave laser at 515 nm operated at 100 µW at the sample. The fluorescent acceptor molecules are excited by a continuous wave laser, at 638 nm operated at 230 µW at the sample. The laser light is guided into the custom-built microscope body and focused by an oil immersion objective (UPLSAPO 60× NA = 1.35, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 20 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (NC395323 - T640lpxr, Chroma, USA). Fluorescence emission was filtered (donor: FF01-571/72-25, acceptor: FF01-679/41-25, Semrock, USA), focused on avalanche photodiodes (SPCM-AQRH-14, Excelitas). The detector outputs were recorded by a national instruments card (PCIe-6353), with acquisition controlled using custom software available on our github (see smfBox reference above).

Data analysis was performed using the PAM software package as described[8] .Single-molecule events were identified using a dual channel burst search algorithm with a threshold of 5 photons, a time window of 500 μs and a minimum photon number of 50.

Lab#8

All sample solutions were measured as a drop on a preliminary passivated with BSA (1mg/ml) coverslip with concentration from 20 to 50pM. The general scheme of the setup which combines pulsed interleaved excitation combined with multiparameter fluorescence detection (PIE-MFD) is described by [4]. In short: Single-molecule FRET experiments with PIE –MFD setup custom built confocal setup which uses time-correlated single photon counting as described previously [60]. The fluorescent donor molecules are excited by a pulsed diode laser (LDH-P-FA 530L, PicoQuant), at 531 nm operated at 20 MHz, 60 μW at the sample in PIE experiment. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-D-C 640, PicoQuant), at 640 nm operated at 20 MHz, 40 μW at the sample in PIE experiment. The laser light is coupled into a fiber collimator (60FC-4-RGBV11-47, Schäfter + Kirchhoff GmbH, Hamburg). And then reflected by a dichroic mirror (HC quadband laser beamsplitter R405/488/532/635, AHF Analysentechnik AG, Tübingen) and focused into the sample by a 1.2 NA water immersion objective (CFI Plan Apochromat VC 60x, Nikon GmbH, Düsseldorf). A diode laser driver (Sepia II, PicoQuant) operates the lasers such that they are pulsed (20 MHz) and shifted (by ~25ns) with respect to each other. The emitted fluorescence is collected by the objective and spatially filtered using a pinhole with 75μm diameter and spectrally split into two beams, parallel and pendicular with respect to the excitation light by a polarizing beam splitter cube (PBS201, Thorlabs, Munich). After the PBS the two beams are then split into donor and acceptor channel according to wavelength (HC BS 649, AHF Analysentechnik) resulting in 2 beams per polarization (green and red). Fluorescence emission was filtered (donor: 582/75 Brightline HC, acceptor: 700/75 ET bandpass, AHF Analysentechnik), focused on a single-photon avalanche diode (τ-SPAD-100, PicoQuant) by a lens of 100 mm focal length. The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant) which is synchronized with the laser driver. The photon arrival times were recorded with 16ps resolution for microtime and synchronization period 50ns for macrotime.

Data analysis was performed using (PAM–PIE analysis with MATLAB v2.0 (develop branch up to commit 320364c4)) software package as described in [8]. Single-molecule events were identified using APBS as burst search method with a threshold of 100 photons per burst, a time window of 500 μs and a minimum photon number of 5. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter was applied using an upper threshold of 10 [9].

Lab#9

Single-molecule fluorescence energy transfer (smFRET) experiments were performed on a custom-built confocal detection-based microscope. The general scheme of the setup was previously described[61] All sample solutions were measured in 8-well chamber slides with a final volume of 200 μl at sample concentration of 50 pM. A pulsed laser diode (LDH-TA-560, Picoquant, Germany), pulsed at 40 MHz, 35 μW, is used for donor excitation; A pulsed laser diode (LDH-D-C-660, Picoquant, Germany), pulsed at 40 MHz, 100 μW, is used for acceptor excitation. We used pulsed interleaved excitation (PIE) scheme to alternatingly excite donor and acceptor fluorophores to retrieve the stoichiometry (S) information[62] .The lasers were cleaned up with excitation filters (Brightline FF01-572/15 and Brightline FF01-650/13, Semrock), passed through the polarizer (GL-10A, Thorlabs) before entering the objective lens (60×, water immersion, NA=1.27, Nikon) through the central dichroic mirror (ZT405/488/561/660/905rpc-UF3, Chroma). The fluorescence emission signal was collected through the same objective lens and spatially filtered using a pinhole with 100 μm diameter and split into parallel and perpendicular polarization axis.

Fluorescence emission was then filtered (donor: Brightline HC612/69, acceptor: ET700/75, Chroma) after separated into donor and acceptor channels via the dichroic mirror (FF650-DI01, Semrock). Photon signals were detected on photon counting detectors (τ-SPAD, Picoquant, Berlin) and recorded by a TCSPC module (Hydraharp400, Picoquant, Berlin). Acquired data were subject to multi-parameter fluorescence analysis and processed burst-wise for fluorescence intensities and lifetime with a threshold of 30 photons per burst[5,22]. All acquired data was performed by a custom written Igor-Program (Wavemetrics)[61] and the burst variance analysis (BVA) was analyzed by using algorithm from PAM-PIE with MATLAB[8] .

Lab#10

All sample solutions were measured as drops on a coverslip with a concentration of 50 pM. A custom-built confocal microscope was used for single-molecule FRET experiments as previously described[63-67] and the setup was modified to allow alternating-laser excitation of donor and acceptor fluorophores[68,69]. For this purpose, the fibre-coupled outputs of a 532-nm laser (operated at 250 μW) and a 640-nm laser (operated at 60 μW) were alternated with a modulation frequency of 20 kHz. Both beams were spatially filtered and coupled into an oil-immersion objective (60x 1.35 NA, UPLSAPO 60XO, Olympus). The same objective was used to collect the resulting fluorescence; the emission was separated from excitation light by a dichroic mirror, focused onto a 200-μm pinhole, and subsequently split spectrally on two avalanche photodiodes (SPCM-AQR-14, PerkinElmer, UK) detecting the donor and acceptor fluorescence with two distinct spectral filters (green: 585DF70; red: 650LP). Custom-written LabVIEW software was used to register and evaluate the detected signal. Fluorescence photons were assigned to either donor or acceptor-based excitation with respect to their photon arrival time. Two characteristic ratios, the fluorophore stoichiometry S and apparent FRET efficiency E*, were calculated for each fluorescent burst, yielding a two-dimensional histogram 6-7. One-dimensional E* distributions for donor-acceptor species were obtained using a $0.4 < S < 0.8$ threshold. These E* distributions were fitted with Gaussian functions, yielding the mean E* value for each distribution.

Lab#11

All sample solutions were measured in glass chambers with concentration 15pM. The general scheme of the setup is described by Krainer et al. 2018[70].In short: Single-molecule FRET experiments with MFD were performed on a homebuilt confocal microscope as described previously by Hartmann et al. 2015[71] .The fluorescent donor and acceptor labelled molecules were excited in PIE mode with 25 MHz repetition rate by a green pulsed diode laser (LDH-P-FA 530L, PicoQuant, Berlin, Germany), at 530 nm wavelength, and a red pulsed diode laser (LDH-D-C 640, PicoQuant), at 640 nm wavelength with a laser power of 110 μW and 90 μW after objective, respectively. The laser light is guided into the inverted confocal microscope (Nikon Eclipse Ti - Nikon, Tokyo, Japan) by a dual-edge beam splitter zt532/640rpc (Chroma, Bellows Falls, VT, USA) and focused by a water immersion objective (FI Plan Apo WI 60x (NA 1.2), Nikon). The emitted fluorescence is collected through the same objective and spatially filtered using a pinhole with 50 μm diameter. In order to detect the fluorescence anisotropy, the spatially filtered light is separated according to its polarization by a polarizing beam splitter (PBS201, Thorlabs, Newton, NJ, USA) and guided in the parallel and perpendicular detection channel. In each detection channel the light is spectrally split into donor and acceptor emission by a single-edge dichroic mirror (FF650-Di01, Semrock, New York, NY, USA). The polarization separated and spectrally split fluorescence emission was bandpass filtered (donor: FF01-582/75, Semrock, acceptor: ET700/75M, Chroma) and focused on avalanche photodiodes (SPCM-AQRH-

14, Excelitas, Waltham, MS, USA). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant).

Data analysis was performed using a custom written software package as described in Hartmann et al. 2017[72] .Single-molecule events were identified from the acquired photon stream as fluorescence bursts with a maximum inter-photon time of 50 µs containing a minimum total number of 40 photons for protein 1 and 100 photons for protein 2 after background correction and a Lee filter with a window size of four. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter as described by Tomov et al. 2012[9] was applied using an upper threshold of 8 (ALEX-2CDE<8).

Lab#12

All sample solutions were measured in Cellview slides (Greiner BioOne). The general scheme of the setup is described[73] .In short: Single-molecule FRET experiments with PIE were performed on commercial PicoQuant MicroTime 200 confocal microscope as described previously [Gouge et al. 2018]. The fluorescent donor molecules are excited by a pulsed diode laser (LDH-P-FA 530, PicoQuant), at 532 nm operated at 40 MHz, 60 µW at the sample in PIE experiment. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-P 640, PicoQuant; clean up: zet636/20x, Chroma), at 640 nm operated at 40 MHz, 30 µW at the sample in PIE experiment. The laser light is guided into the epi-illuminated confocal microscope (Olympus IX73, Hamburg, Germany) by dual-edge beam splitter ZT532/640rpc-UF3 (Chroma) focused by a water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (T635lpxr; Chroma). Fluorescence emission was filtered (donor: ff01-582/64; Semrock, acceptor: H690/70; Chroma), focused on avalanche photodiodes (SPCM-AQRH-14-TR, Excelitas). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant).

Data analysis was performed using the PAM software package as described[8]. Single-molecule events were identified using an APBS and DCBS with a threshold of 100 counts, a time window of 500 µs and a minimum photon number of 10 (both channels for DCBS). Histograms were corrected for donor leakage and acceptor direct excitation, γ and β.

Lab#13

All sample solutions were measured as a drop on a coverslip (Roth, Karlsruhe, burned at 500 °C for 2 h), covered with a humidity chamber with a dilution of 1 to 300 (sample 1) and 1 to 200 (sample 2 and 3) of the delivered stock solution. The general scheme of the setup is described[74]. In short: Single-molecule FRET experiments with PIE were performed on a homebuilt confocal microscope as described[74] .The fluorescent donor molecules are excited by a cw DPSS “Crysta Laser” (GCL-005-L, Laser2000, Wessling) at 532 nm with 40 µW at the sample in the PIE experiment. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-P-C 635, PicoQuant GmbH), at 635 nm operated at 10 MHz, 5 µW at the sample in PIE experiment. The laser light is guided through a single-mode fibre (SMC-460, Schäfter&Kirchoff), directed by dual-band beamsplitter (Z532/633, Chroma, Bellows Falls, USA) and then focussed by a water immersion objective (CFI PlanApochromat 60x WI, Nikon, Japan). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (BS640DCXR, Chroma, Bellows Falls, USA). Fluorescence emission was filtered (donor: FF01_582_75, Semrock, USA) acceptor: HQ685_70, AHF Analysentechnik AG, Tübingen, Deutschland), focused on avalanche photodiodes (SPCM-AG 14, Perkin

Elmer, Fremont, USA). The detector outputs were recorded by a TCSPC module (TimeHarp200, PicoQuant GmbH).

Data analysis was performed using IgorPro 8 (Wavemetrics). Single-molecule events were identified using a bin-selection algorithm with a threshold of 40 photons in the sum of donor and acceptor channel upon donor excitation and a threshold of 10 photons in the acceptor channel upon acceptor excitation.

Lab#14

Our multi-parameter fluorescence detection setup equipped with pulsed interleaved excitation is conceptually identical to the confocal microscope described previously[4]. For excitation, a pulsed supercontinuum laser was used with wavelength selector at 532±5 nm (Solea, Picoquant, Berlin, Germany), and a spectrally filtered (Chroma z635/10x, Picoquant) 635-nm laser diode (LDH-P-C-635B, Picoquant). Both lasers were alternated at 26.67 MHz (PDL 828 Sepia2, Picoquant), delayed ~18-ns with respect to each other and combined via a dichroic mirror (Chroma T560lpxr, F48-559, AHF) in a single-mode optical fiber (coupler: 60FC-4-RGBV11-47, fiber: PMC-400Si-2.6-NA012-3-APC-150-P, Schäfter und Kirchhoff GmbH, Hamburg, Germany). After collimation (60FC-L-4-RGBV11-47, SuK GmbH), the linear polarization was cleaned up (Codixx VIS-600-BC-W01, F22-601, AHF) and the light was reflected on a 3-mm thick excitation polychroic mirror (Laser Beamsplitter zt532/640rpc, F58-PQ09, AHF) upward and into the back port of the microscope (IX70, Olympus Belgium NV, Berchem, Belgium) via two mirrors and upward to the sample (3-mm thick Full Reflective Ag Mirror, F21-005, AHF, mounted in a TIRF Filter Cube for BX2/IX2, F91-960, AHF) to the objective (UPLSAPO-60XW, Olympus). Sample emission was focused through a 75-µm pinhole (P75S, Thorlabs, Munich, Germany) via an achromatic lens (AC254-200-A-ML, Thorlabs), collimated again (AC254-50-A-ML, Thorlabs) and spectrally split (Laser Beamsplitter H 643 LPXR superflat, F48-643, AHF). The yellow range was filtered (582/75 BrightLine HC, F37-582, AHF) and polarization was split (PBS251, Thorlabs). The red range was also filtered (Chroma ET705/100m, AHF) and polarization was split (PBS252, Thorlabs). Photons were detected on four avalanche photodiodes (yellow photons: Laser Components COUNT Blue, red photons: Perkin-Elmer or EG&G SPCM-AQR12/14) all of which were connected to a time-correlated single photon counting (TCSPC) device (SPC-630, Becker & Hickl GmbH, Berlin, Germany) over a router (HRT-82, Becker & Hickl) and power supply (DSN 102, Picoquant). Signals were stored in 12-bit first-in- first-out (FIFO) files. All analyses of experimental data were performed in the software package PAM[8].

Lab#15

All sample solutions were measured in Corning 384 well non-binding plates at concentrations ranging between 20-60 pM. The general scheme of the setup is described[75].In short: Single-molecule FRET experiments with PIE - MFD were performed on a homebuilt confocal microscope as described previously[75].Excitation was performed with a pulsed SC450-4-20MhZ supercontinuum source (Fianium, Southampton, UK). It runs at 20MHz, and has a power density >2mW/nm over the 450-800nm range, with an average pulse width of 100-150ps. The collimated, unpolarized output of the source is divided by a 50:50 beamsplitter cube (BS016, Thorlabs, NJ, USA), thus generating two beams. The beams are spectrally filtered using excitation bandpass filters at wavelength 532/10 (prompt beam) and 635/10 (delayed beam) to excite donor and acceptor molecules, respectively. The delayed beam has a path length of ~ 8m relative to the prompt beam, generating a ~24ns delay in the pulse. The two beam paths are then recombined using a 50:50 beamsplitter cube (BS016, Thorlabs, NJ, USA) and focused using a 10x objective into a single-mode fiber (SMF) (P1-460A-FC, Thorlabs, NJ, USA), by which the beams become spatially overlapped and filtered. The output of the fiber is collimated using a 10x microscope objective lens (04OAS010; CVI Melles Griot, Albuquerque, NM, USA), polarized by a polarizing beamsplitter cube (PBS; PBS201,

Thorlabs, NJ, USA) and coupled into an inverted microscope (Eclipse Ti, Nikon, France). Power was controlled using ½ (prompt: WPH05M-532 and delayed WPH05M-633, Thorlabs, NJ, USA) and ¼ waveplates (prompt: WPQ05M-532 and delayed WPQ05M-633, Thorlabs, NJ, USA) placed in the prompt and delayed beams before recombination, resulting in 50 µW for the prompt and 25 µW for the delayed beam at the entrance into the microscope. The light is reflected by a dichroic mirror that matches the excitation/emission wavelengths (FF545/650-Di01, Semrock, Rochester, NY, USA) and coupled into a Nikon 100x, NA1.4 objective. Emitted photons are then collected by the objective and focused by the tube lens on a pinhole of 150µm. The emission photon stream is collimated and divided using a polarizing beamsplitter cube (PBS; PBS201, Thorlabs, NJ, USA). In each created polarization channel, the photons are spectrally separated using dichroic mirrors (BS 649, Semrock, Rochester, NY, USA) and filtered using high quality emission filters (parallel: ET BP 585/65 and ET BP 700/75 and perpendicular: HQ 590/75 M and HQ 660 LP, Chroma, Bellows Falls, VT, USA). Single photons are detected using Single Photon Avalanche Diodes. We use two MPD-1CTC (MPD, Bolzano, Italy) for the donor wavelength channels and two SPCM AQR-14 (Perkin Elmer, Fremont, CA, USA) for the acceptor wavelength channels. The output of the detectors is coupled into a TCSPC counting board (SPC-150, Becker&Hickl, Berlin, Germany), through a HRT41 router (B&H), using appropriate pulse inverters and attenuators. The sync signal that triggers the TCSPC board is provided by picking a small fraction of the light from the prompt path (reflected by a coverslip), and focusing it on an avalanche diode (APM-400, B&H).

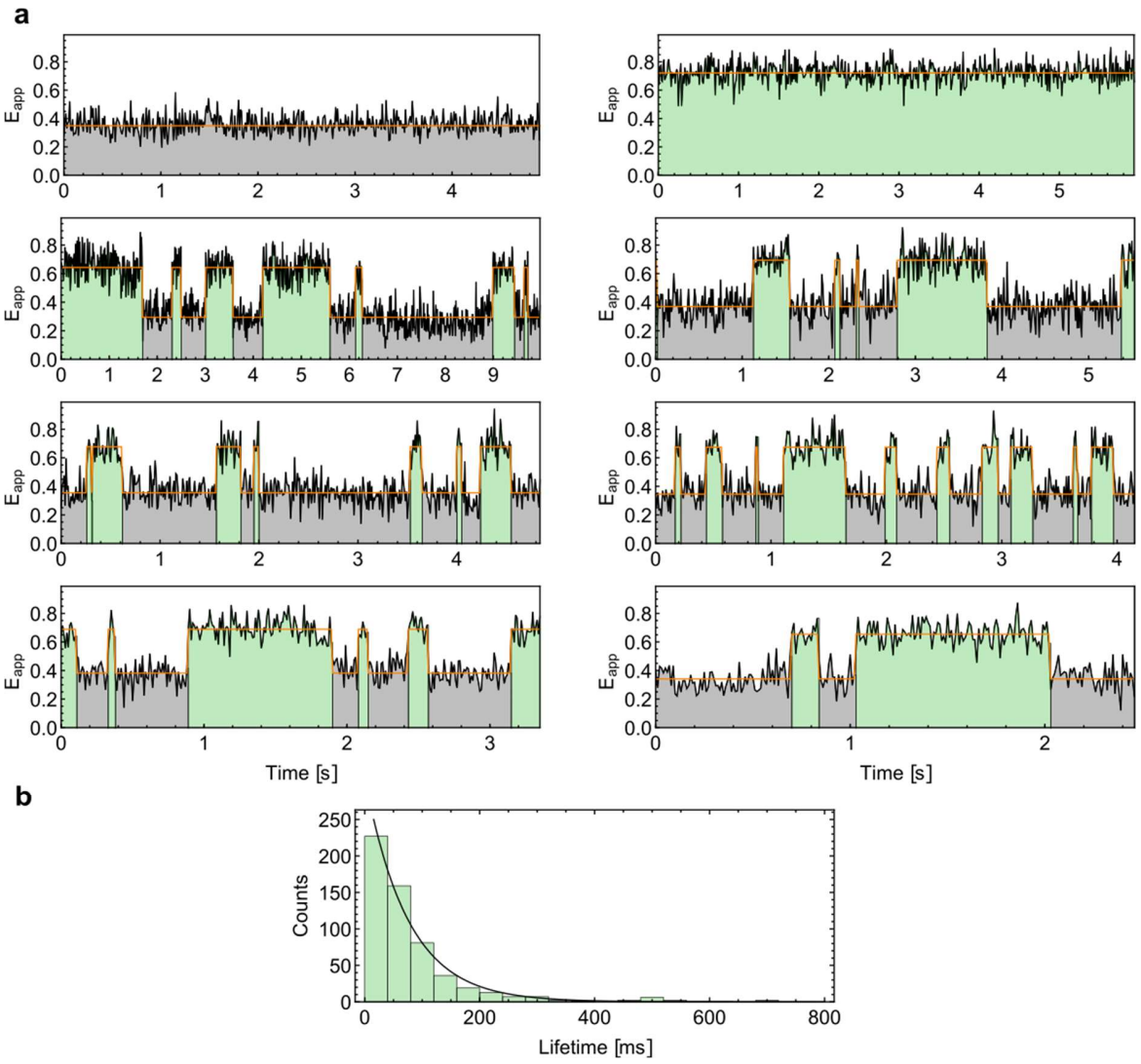
Data analysis was performed using the Software Package for Multiparameter Fluorescence Spectroscopy, Full Correlation and Multiparameter Fluorescence Imaging developed in C.A.M. Seidel's lab (http:// www.mpc.uni-duesseldorf.de/seidel/). A single-molecule event was defined as a burst containing of at least 40 photons with a maximum allowed interphoton time of 0.3 ms and a Lee-filter of 20. Photobleaching events were identified base on |TGX-TRR|<1 ms as described [4].

Lab#16

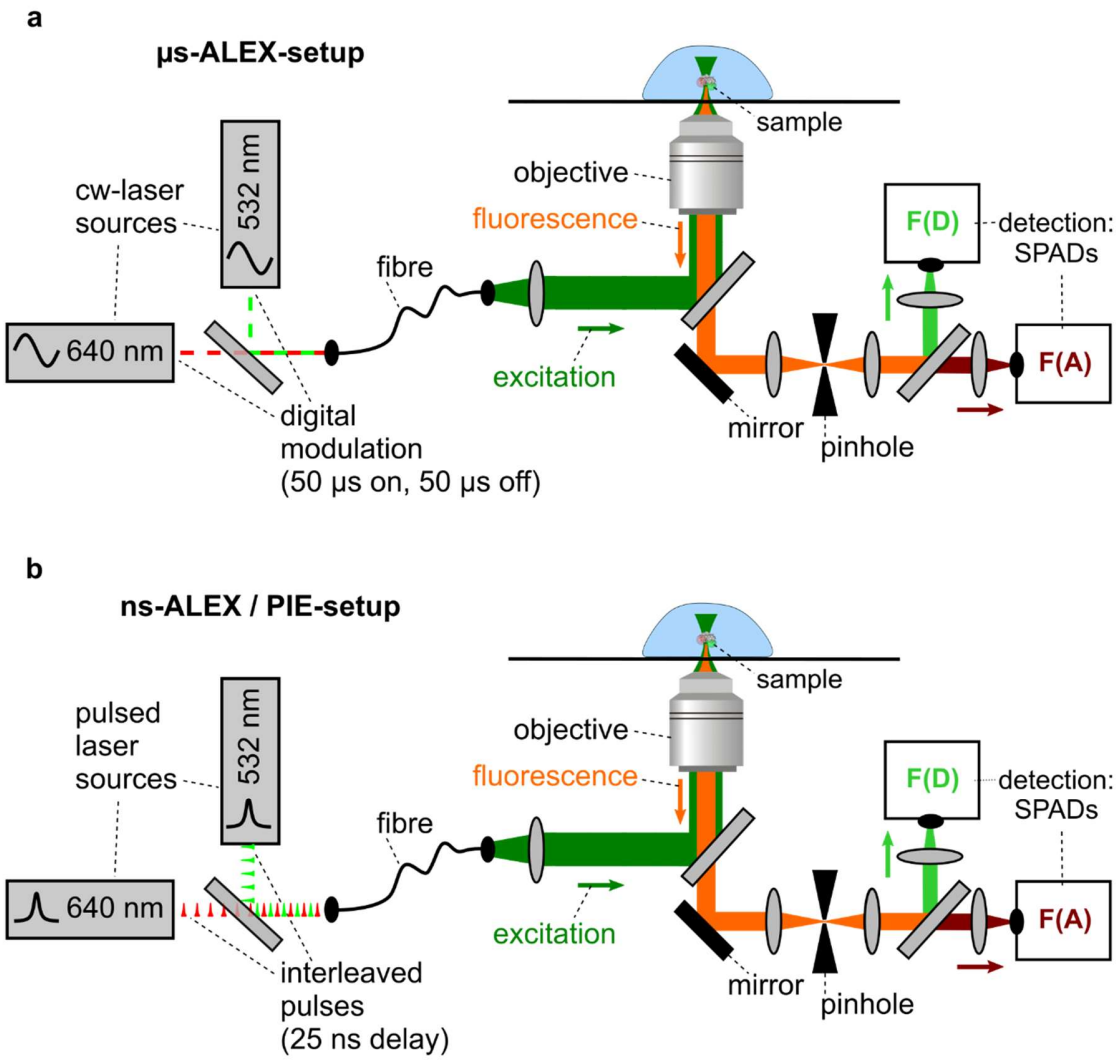
All sample solutions were measured in chamber with concentration 100 pM. The general scheme of the setup is similar to Supplementary Figure 2. In short: Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal microscope as described previously. The light from 532nm (Compass 215M-50, Coherent, USA) and 638nm (25mW Red flame, Coherent, USA) cw-laser sources is alternately directed to an IX71 inverted microscope (Olympus, Japan) every 25 µs, reflected on a beamsplitter (Z488-533-633RPC, Chroma, USA) and focused through a water-immersion objective (UPlanApo 60x/1.2w, Olympus, Japan) to excite fluorescent molecules. The light intensities were 100 µW for 532nm and 35 µW for 638nm, measured before the beamsplitter. The emitted fluorescence is collected through the objective, spatially filtered using a 100 µm pinhole, and then spectrally split into two photon streams by a dichroic mirror (635 DCXR, Chroma, USA). Individual photon streams were filtered (for donor: HQ580/60m, for acceptor: HQ665lp, Chroma, USA) and detected by single photon-avalanche photodiodes (SPCM-AQR-14, PerkinElmer, USA). The detector outputs (photon arrival times) were recorded by a counter/timer device module (PCI-6602, National Instruments, USA).

Data analysis was performed using the ALEX-suite software package as described[76].Single-molecule events were identified using an all-photon-burst-search (APBS) and a dual-channel-burst-search (DCBS) algorithm with a threshold of 150, a time window of 50 µs and a minimum photon number of 50.

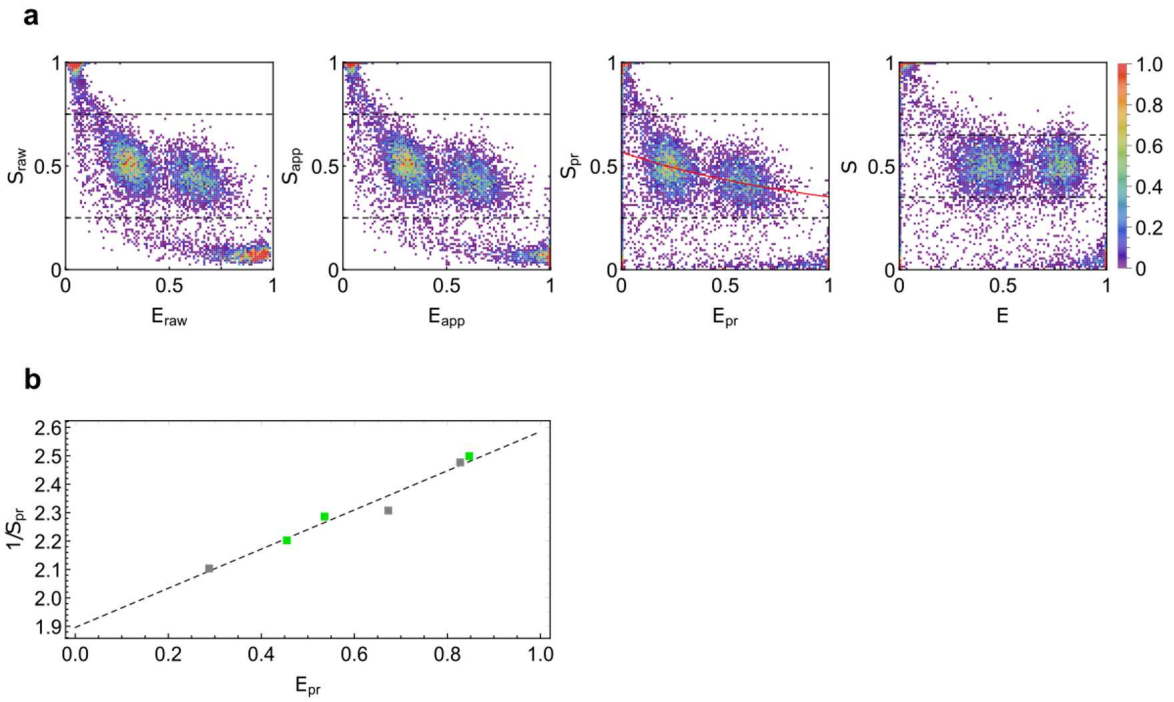
Supplementary Figures



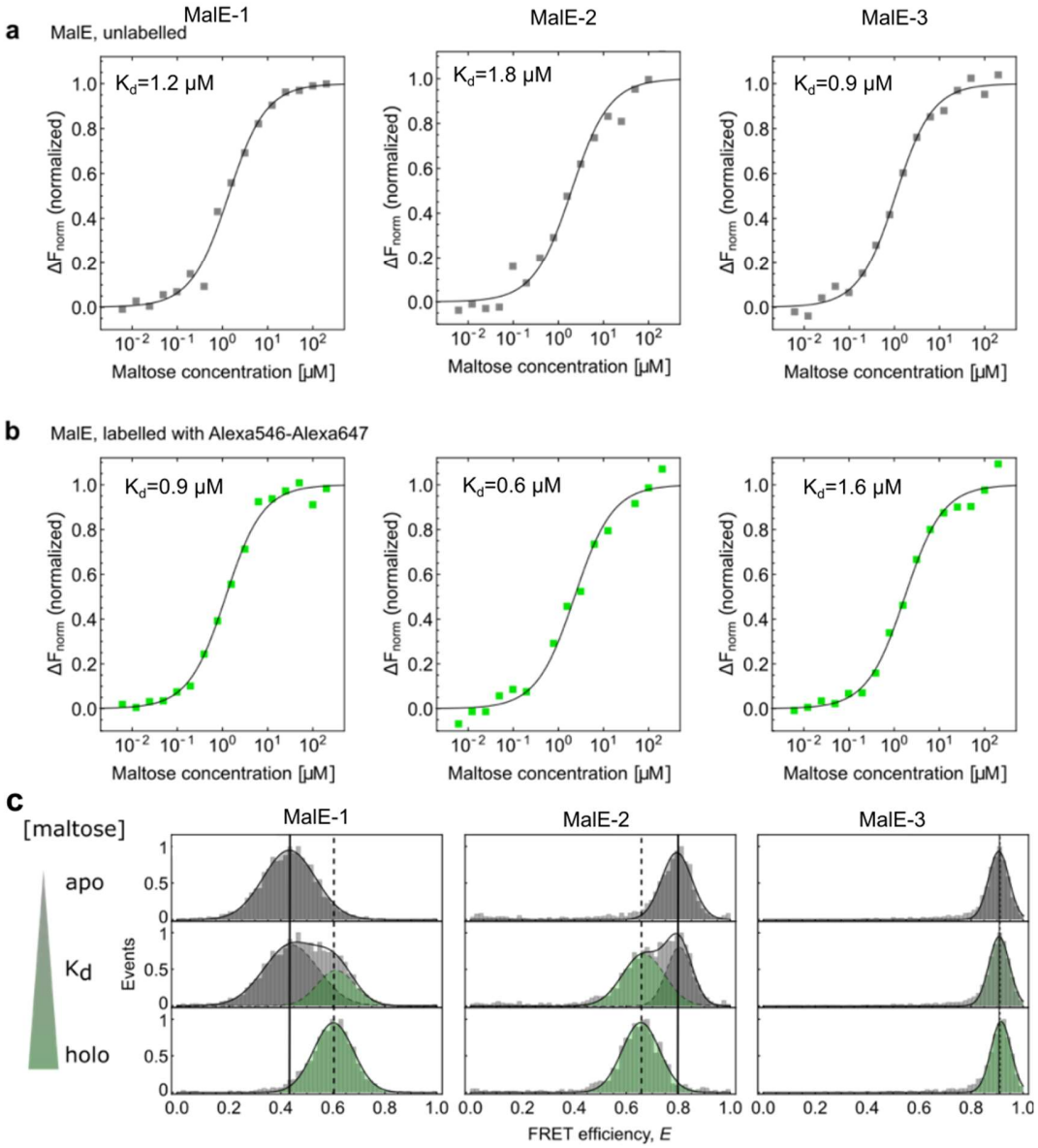
Supplementary Figure 1. Ligand-induced slow conformational dynamics of MalE switching between the apo- and holo states. **(a)** Time traces of immobilized MalE 36C/352C molecules labeled with Alexa555-Alexa647 measured in buffer containing no ligand (top row, left, grey), 1 mM maltose (top row, right, green) and 1 µM maltose , close to the K_d (rows 2-4)[77]. The samples were measured in a scanning confocal microscope as described in reference 7. FRET states and lifetimes are extracted from a fitted two-state Hidden-Markov-Model as described in reference 78. The traces show ligand-induced interconversion of states on the >10 millisecond timescale. **(b)** A dwell time histogram of the duration of the holo state with an exponential fit (solid line) shows a mean dwell time of 75 ms.



Supplementary Figure 2. Schematics of the experimental setups. **(a)** A schematic of a confocal microscope setup used for the acquisition of diffusion-based smFRET data using alternating-laser excitation (ALEX). Continuous wave laser sources (here with wavelengths of 532 nm and 640 nm) excite the sample alternatively for periods of 50 µs through a microscope objective. F(D) and F(A) indicate the donor and acceptor detection channels, respectively. **(b)** A schematic of a confocal microscope setup used for the acquisition of diffusion-based smFRET data using nsALEX / pulsed interleaved excitation (PIE). Pulsed laser sources (here with wavelengths of 532 nm and 640 nm) alternatingly excite the sample with picosecond pulses delayed by ~25 ns.



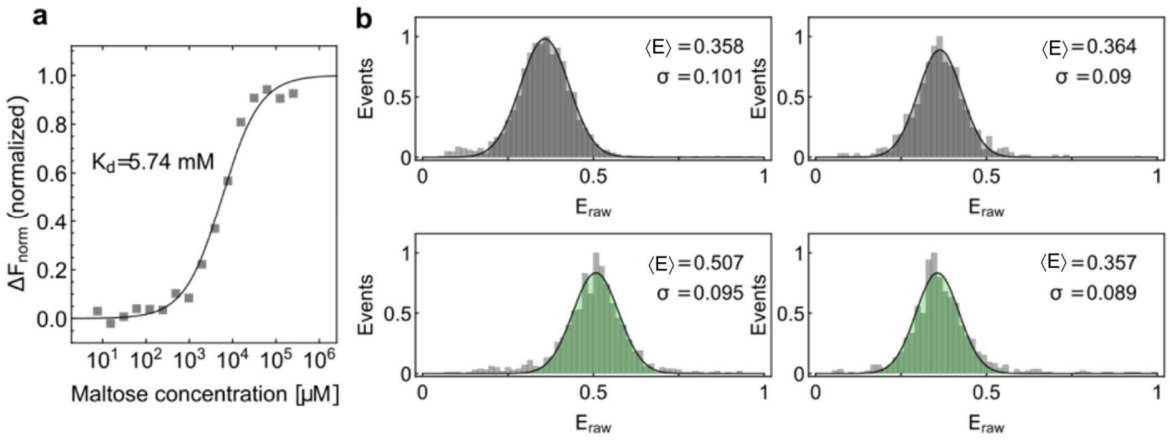
Supplementary Figure 3. Correction procedure when using a global γ. **(a)** *ES*-diagrams of two selected and merged data sets (Alexa546-Alexa647 labeled MalE-1 apo and MalE-2 apo) showing the results from the different correction steps; from left to right: raw data, background corrected apparent FRET efficiency, crosstalk and direct excitation corrected proximity ratio, E_{pr}, with fitted γ curve (red), and the γ corrected FRET efficiency versus stoichiometry plots. **(b)** Proximity ratio, E_{pr} and stoichiometry S_{pr} of all MalE mutants in the apo (gray dots) and holo states (green dots) with the linear fit (dashed line) used for a global γ correction.



Supplementary Figure 4. Binding affinity measurements of maltose to MalE using microscale thermophoresis. **(a)** The binding affinities of maltose to MalE were measured using microscale thermophoresis (Monolith NT.LabelFree, Nanotemper) where the ratio of fluorescence before and after heating $\Delta F_{norm} = F_{cold}/F_{ho}$ was recorded at different maltose concentrations[79]. Data points show ΔF_{norm} normalized to the minimal and maximal fluorescence intensities for the unlabeled mutants MalE-1 (left), MalE-2 (middle), and MalE-3 (right). The curves were fitted with a standard model for receptor-ligand kinetics

$$\Delta F_{norm} = \frac{K_d + c_P + c_{malt} - \sqrt{(K_d + c_P + c_{malt})^2 - 4 c_P\, c_{malt}}}{2\, c_P},$$

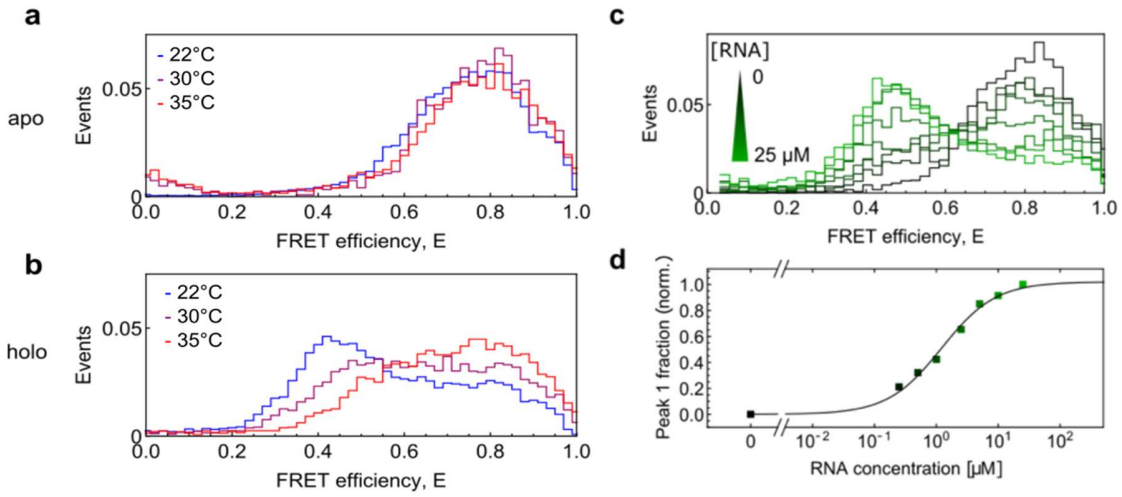
where K_d is the dissociation constant, c_P the protein concentration (0.25 µM in the experiment), and c_{malt} the maltose concentration. The fits to the binding model (solid line) yield K_d-values of 1.2 µM (left), 1.8 µM (middle), and 0.9 µM (right), respectively. **(b)** The binding affinities of maltose to fluorescently-labeled MalE (Alexa547 and Alexa647) mutants MalE-1 (left), MalE-2 (middle), and MalE-3 (right) were measured using microscale thermophoresis. These experiments yielded K_d-values of 0.9 µM (left), 0.6 µM (middle), and 1.6 µM (right), respectively. **(c)** FRET efficiency E histogram for the MalE mutant MalE-1 (left), the mutant MalE-2 (middle), and the mutant MalE-3 (right) in the presence of 0 (top), 1 µM (middle), and 1 mM maltose (bottom).



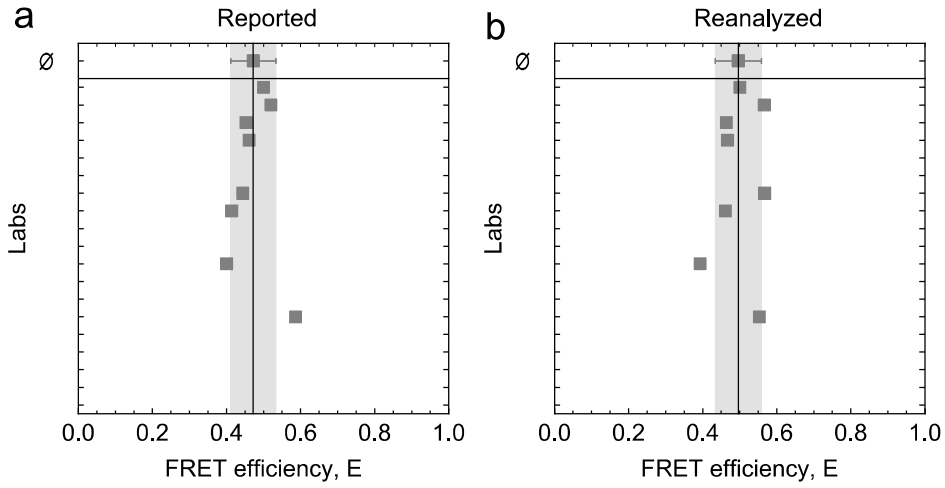
Supplementary Figure 5. Millimolar maltose concentration does not influence fluorophore properties. To investigate the potential influence of maltose on the photophysical properties of the used fluorophores, measurements were performed using the D65A closing deficient mutant of MalE-1. **(a)** The binding affinity of maltose to the MalE-1 mutant D65A were measured using microscale thermophoresis. The ratio of fluorescence before and after heating $\Delta F_{norm} = {F_{cold}}/{F_{hot}}$ was recorded at different maltose concentrations[79]. Data points show ΔF_{norm} normalized to minimal and maximal fluorescence intensities for the unlabeled mutants. The curves were fitted with a standard model for receptor-ligand kinetics:

$$\Delta F_{norm} = \frac{K_d + c_P + c_{malt} - \sqrt{(K_d + c_P + c_{malt})^2 - 4 c_P\, c_{malt}}}{2\, c_P},$$

where K_d is the dissociation constant, c_P the protein concentration (0.25 µM in the experiment) and c_{malt} the maltose concentration. A K_d of 5.7 mM was measured. **(b)** Raw FRET efficiency E_{raw} histograms for the MalE mutant MalE-1 labeled with Alexa546-Alexa647 without maltose (top, left) and with added 1 mM maltose (bottom, left) compared to same measurement with the MalE-1 closing deficient mutant D65A (right). No influence of the maltose is visible in the FRET efficiency histogram for the closing deficient mutant.

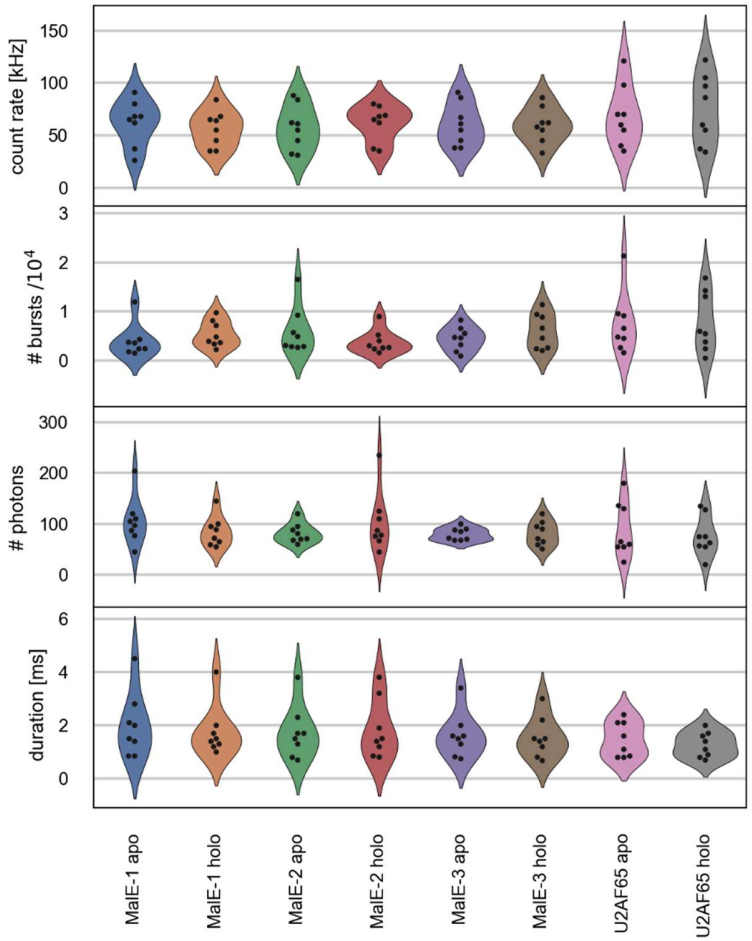


Supplementary Figure 6: Temperature and concentration dependence of RNA binding to U2AF65. **(a, b)** SmFRET histograms of **(a)** apo and **(b)** holo measurements at 22°C (blue), 30°C (purple) and 35°C (red). **(c)** SmFRET histograms for U9 RNA titration measurements with U2AF65 (low to high RNA concentrations are shown in a color gradient from black to light green (0, 0.25, 0.5, 1.0, 2.5, 5, 10 and 25 µM). **(d)** The area under peak 1 of the FRET histograms (0.1-0.6 FRET efficiency) from panel c is plotted versus the U9 RNA concentration to estimate the K_d. For normalization, the area for the apo measurement was set to zero and for the holo measurement at 25 µM was set to 1. The affinity of U9 RNA binding to U2AF62 was estimated to be ~1.2 µM using the standard model for receptor-ligand kinetics as described in Supplementary Figure 4.

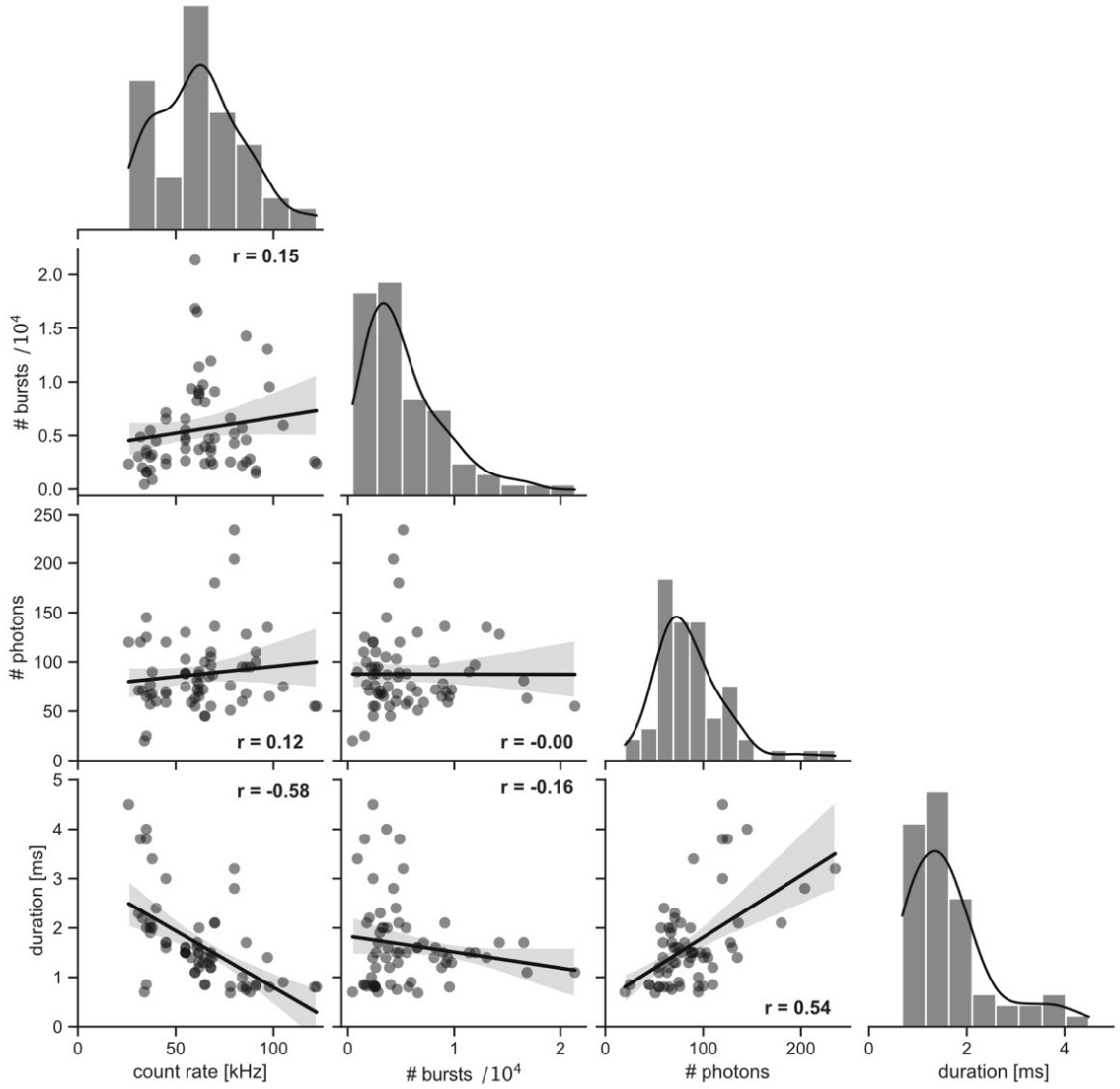


Supplementary Figure 7. Comparison of the FRET efficiency reported for different labs and after reanalysis for the MalE-1 mutant under apo conditions: **(a, b)**. The determined average FRET efficiency values (**a**) reported by the different laboratories and (**b**) after reanalysis are shown as squares for measurements of MalE-1 in the apo states for 8 laboratories. The mean value (upper data point) from all data sets with the corresponding standard deviation is shown in grey. For the details of the reanalysis procedure, please refer to the Supplementary Note 4. The mean FRET Efficiency and standard deviation was 0.472 and 0.061 for the reported values

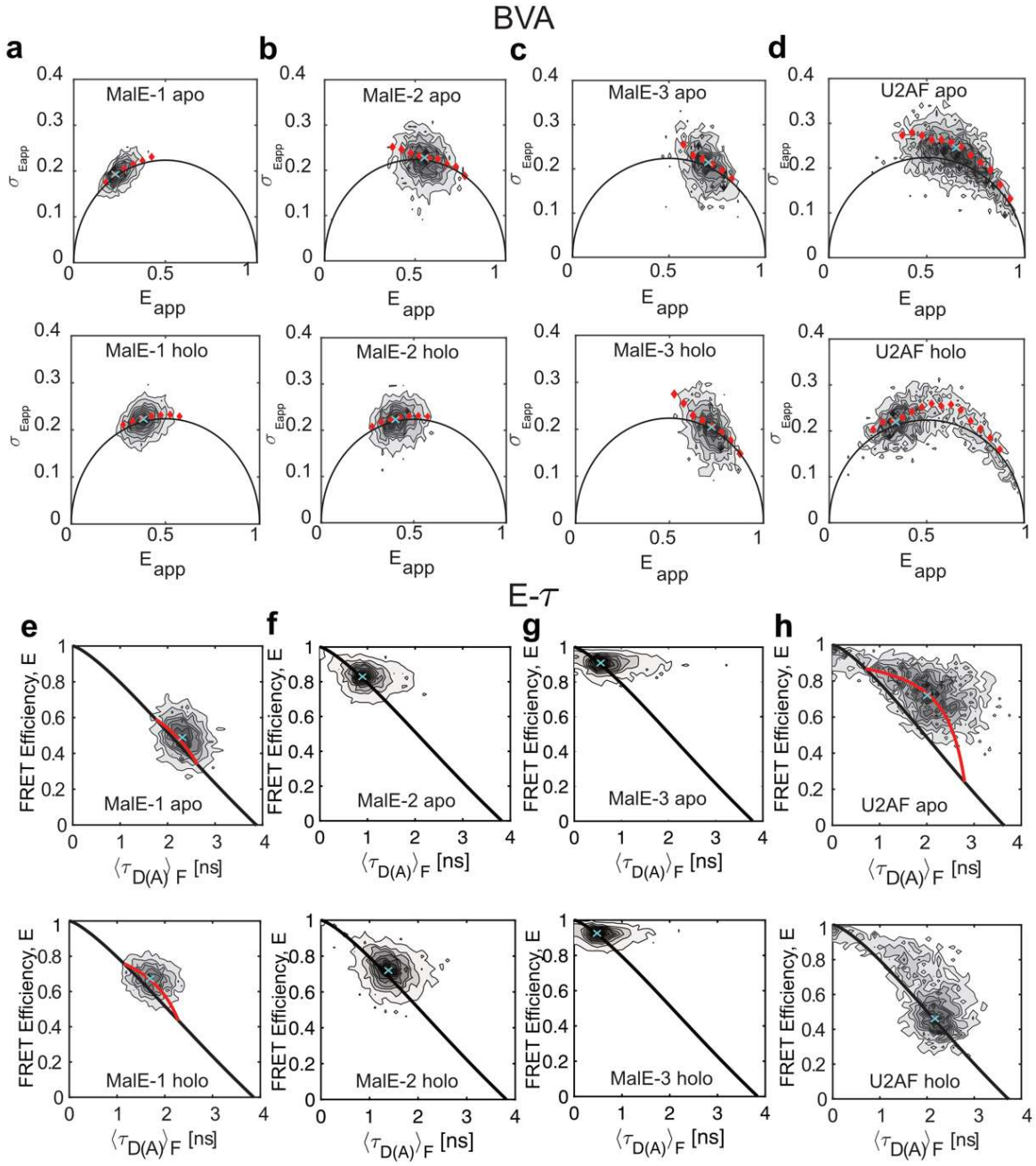
and 0.496 and 0.062 after reanalysis respectively. For reanalysis purposes, we used data from the same 8 labs which measured the dynamics of U2AF.



Supplementary Figure 8: Sample-dependent setup statistics. Sample-dependent distributions of the setup-dependent parameters are given as violin plots for (*top*) the average count rate, (*second from the top*) number of detected bursts, (*third from the top*) average number of photons per burst and (*bottom*) average burst duration. The procedure to obtain the values for the above parameters from the measurement data collected from 8 labs for both MalE and U2AF samples is as follows: The collected raw data from different labs was analyzed using the PAM software[88]. Briefly, first, a burst search was performed using an all photon burst search with a threshold of 50-100 photons per sliding time window of 500 µs depending on the dataset. For one set of measurements, a lower threshold of 20 photons per 500 µs time window was necessary. After burst selection, background subtraction and correction for crosstalk and direct excitation were performed as discussed in the data analysis section. To remove blinking and bleaching events as well as for selecting out the double-labeled molecules, an ALEX-2CDE filter[9] with a lower limit of 5 and an upper limit of 25 was used depending on the data set. The ALEX-2CDE filter values differed depending on the excitation intensities and sample concentrations used for the measurements. The values for all the parameters are a median of the values obtained for the double-labeled molecules for each measurement. These values were made available with PAM software.

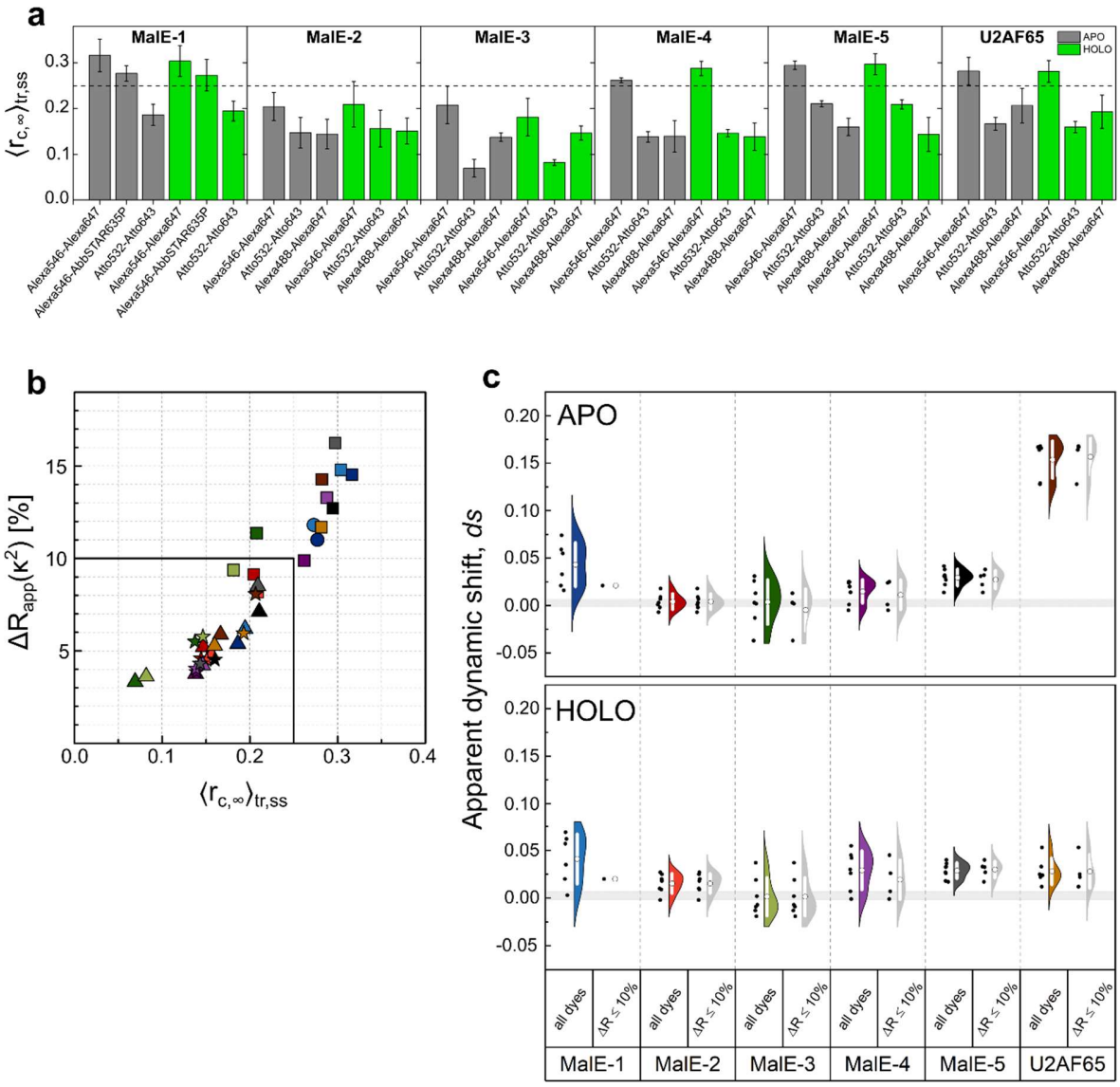


Supplementary Figure 9: Correlations between setup-dependent parameters. Pairwise plots of the number of detected bursts, average number of photons per burst, average burst duration and average burstwise count rate shown in Supplementary Figure 8. *r* is the Pearson's correlation coefficient. The one-dimensional projections show the distribution of the parameters as histograms (gray bars) and kernel density estimates (black lines).

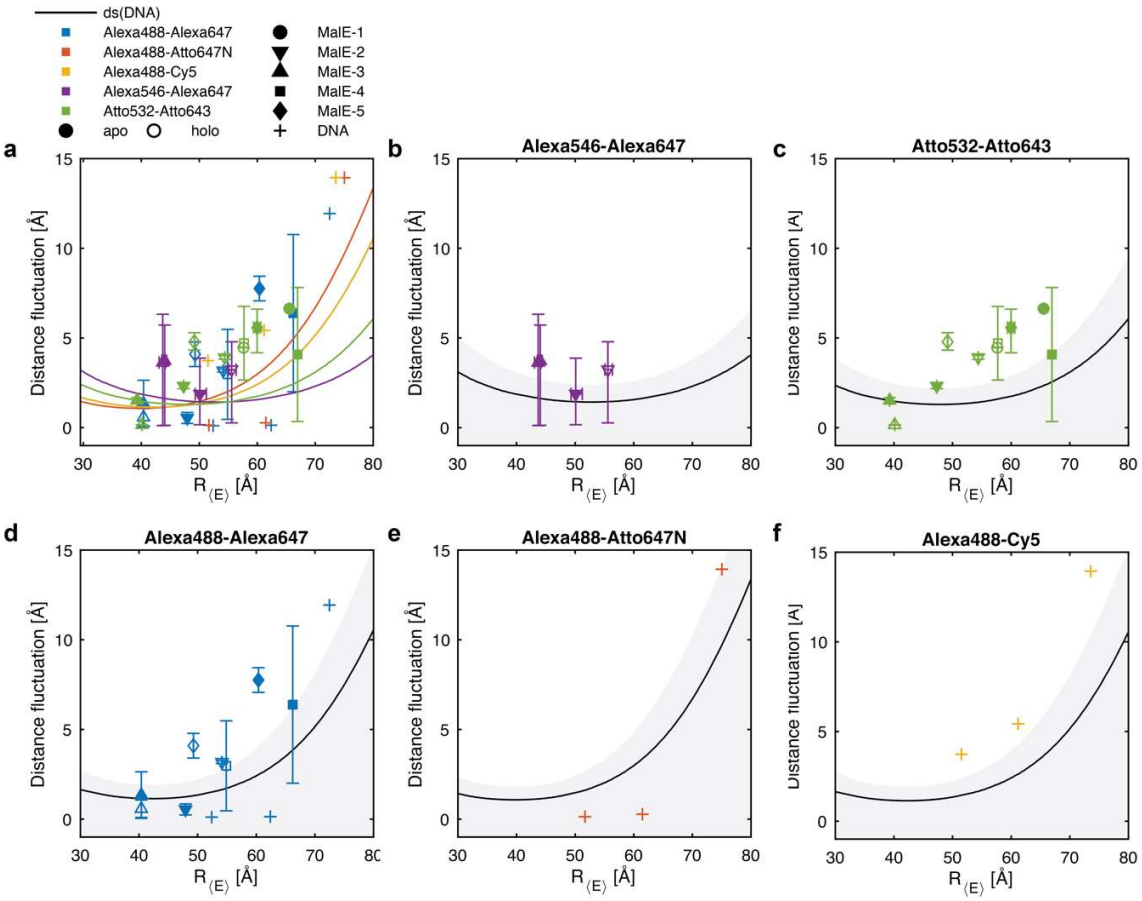


Supplementary Figure 10. Overview of conformational dynamics and determination of the dynamic shift on the sub-millisecond timescale in MalE labeled with Alexa546-Alexa647 and U2AF65 labeled with Atto532-Atto643: **(a-d)** Burst variance analysis (BVA) of MalE-1 (a), MalE-2 (b), MalE-3 (c), and U2AF (d) under both apo (upper panel) and holo (lower panel) conditions. In BVA, the standard deviation $\sigma_{E_{app}}$ of the apparent FRET efficiency E_{app} is compared to the shot-noise limit given by $\sigma_{E_{app}} = \sqrt{E_{app}(1-E_{app})/n}$ (black line, here $n = 5$). Single-molecule events with conformational dynamics show an increased variance and follow the dynamic line (red diamonds). Red diamonds indicate the average standard deviation of all bursts within a FRET efficiency range of 0.05. The mean positions of the populations (cyan crosses) were determined by fitting a two-dimensional Gaussian distribution to the data (see Supplementary Note 6). The dynamic shift, *ds*, is defined as the excess standard deviation compared to the static line. A clear deviation from the static line is observed for the apo state of U2AF65. For U2AF65 under holo conditions, the *ds* was determined for the major holo state

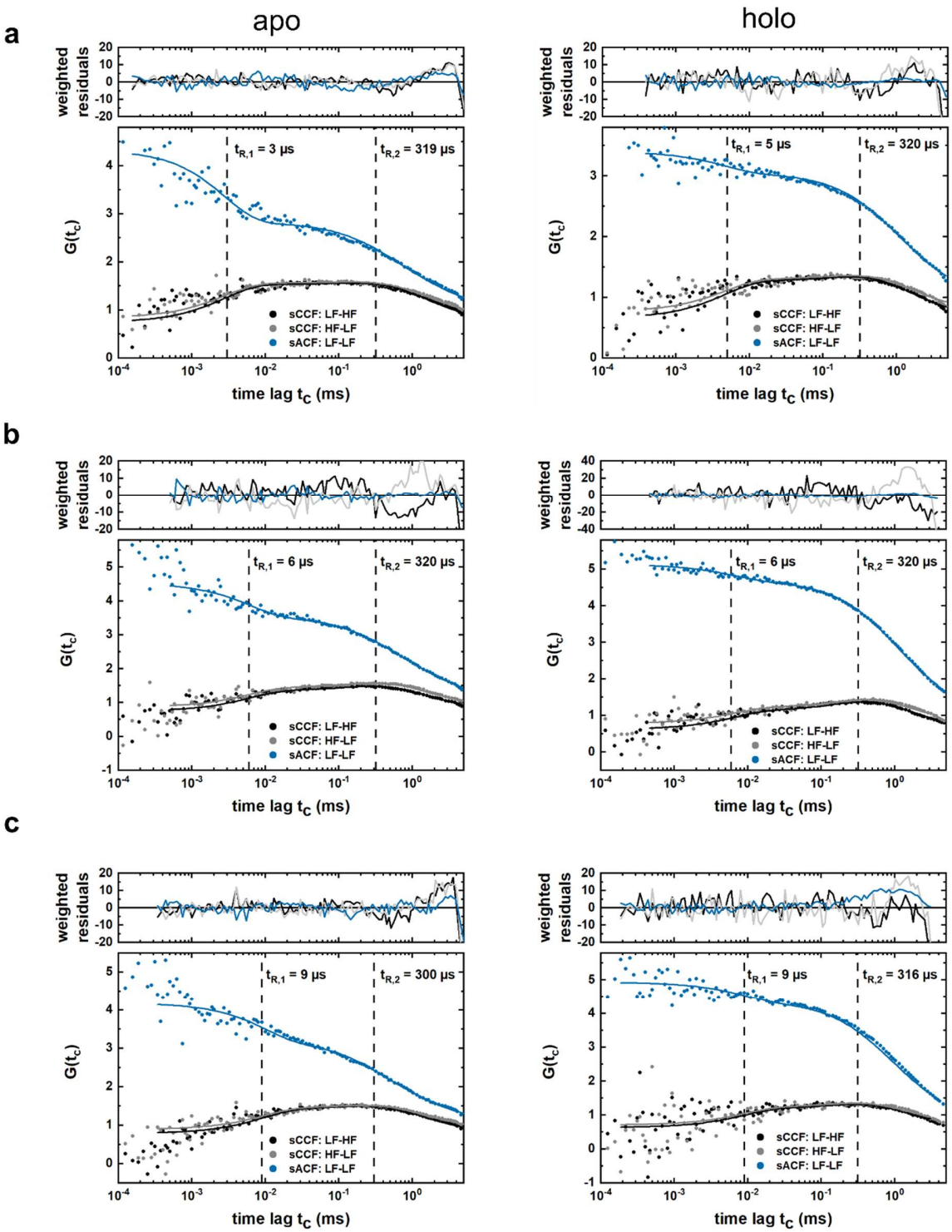
population. Please note the leftover minor apo state population in holo condition has a similar deviation as for the apo condition. **(e-h)** Plots of the FRET efficiency *E* versus intensity-weighted average donor lifetime $\langle\tau_{D(A)}\rangle_F$ (E-τ) for MalE-1 (e), MalE-2 (f), MalE-3 (g), and U2AF (h) under both apo (upper panel) and holo (lower panel) conditions. In the E-τ plot, the intensity-based FRET efficiency *E* is plotted against the intensity-weighted average donor fluorescence lifetime, $\langle\tau_{D(A)}\rangle_F$. The static FRET-line is given by the Förster relation as $E = 1 - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}$ (black). The static lines are slightly curved as they account for the flexibility of the dye linkers. Molecules undergoing dynamics are shifted from the static line and follow a dynamic FRET-line (red). The mean positions of the populations (cyan crosses) were determined by fitting a two-dimensional Gaussian distribution to the data (see Supplementary Note 6). For a given population, the dynamic shift is defined as the displacement of the population orthogonal to the static FRET-line. A clear dynamic shift is observed for U2AF65. The slight dynamic shift observed in MalE-1 is due to the high anisotropy of attached dyes as identified later in the study (see Fig. 5e,f and Supplementary Figure 11). The end points of the dynamic FRET-line for MalE-1 and U2AF65 were determined from a sub-ensemble analysis of the fluorescence decay. For U2AF65 under holo conditions, the *ds* was determined for the major holo state population. Please note that the leftover minor apo state population in the holo measurement has a similar deviation as for the apo condition.



Supplementary Figure 11. Dye-specific sticking interactions in proteins and residual apparent dynamic shift upon removal of sticky dyes. **(a)** Combined residual anisotropies for different dye combinations of MalE and U2AF65 variants (see Supplementary Note 9 and Supplementary Table 12) **(b)** The distance uncertainty due to the orientation factor κ^2, $\Delta R_{\mathrm{app}}(\kappa^2)$ (see Supplementary Note 10 and Supplementary Table 13), and its correlation with the combined residual anisotropy $\langle r_{c,\infty}\rangle_{tr,ss}$. A threshold of 10% in terms of distance uncertainty (solid black line) corresponds to $\langle r_{c,\infty}\rangle_{tr,ss}$= 0.25, which was used to filter-out sticking-prone dyes. **(c)** Violin plots of the dynamic shift values reported by three labs (see Supplementary Table 7) including all measured dye combinations and after filtering of the dyes to remove those with pronounced sticking interactions using a threshold of $\Delta R_{\mathrm{app}}(\kappa^2) = 10\%$ corresponding to $\langle r_{c,\infty}\rangle_{tr,ss}$ = 0.25. In the violins, the white circle and whisker denote the mean ± 1σ. The shaded grey area gives the upper and lower limit of apparent dynamic shift obtained from DNA samples with different FRET efficiencies and with different dye combinations (Supplementary Table 7).

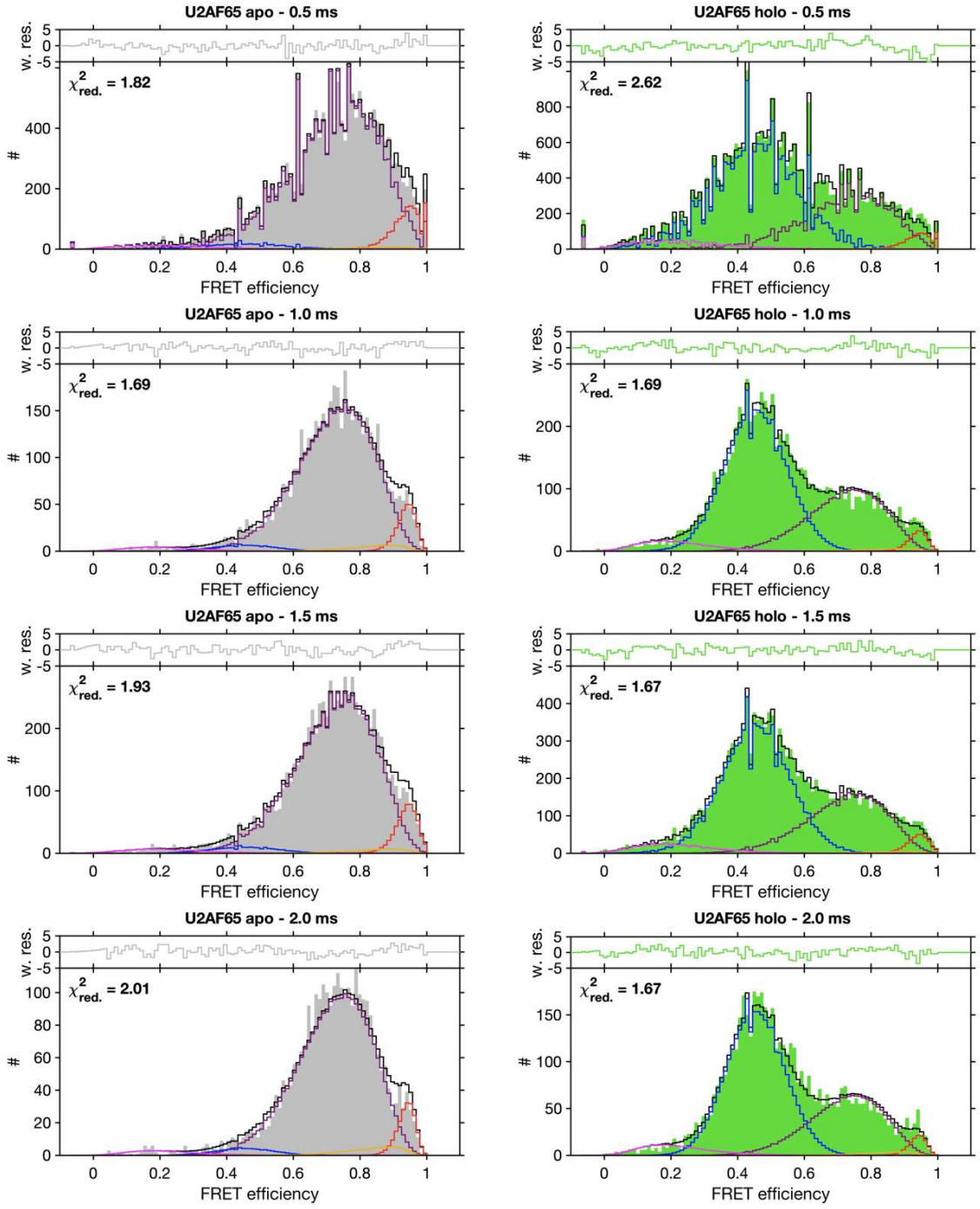


Supplementary Figure 12. Estimated conformational flexibility of MalE based on the residual dynamic shift for different donor-acceptor pairs. **(a)** The estimated distance fluctuations are plotted against the FRET-averaged interdye distance for the give MalE mutants in the apo and holo state labeled with the dye pairs Alexa546-Alexa647, Atto532-Atto643 and Alexa488-Alexa647. Additional control measurements on dsDNA are shown as crosses for the dye pairs Alexa488-Alexa647, Alexa488-Atto647N and Alexa488-Cy5. The lines indicate the apparent distance fluctuation obtained for the dsDNA control measurements, calculated based on a dynamic shift of $ds_{DNA} = 0.0026 \pm 0.0044$. **b-f)** Individual plots of the data shown in panel a for the different dye combinations. Gray areas indicate the 1σ confidence interval for the apparent distance fluctuations obtained for the dsDNA control measurements. All values are given in Supplementary Table 7.



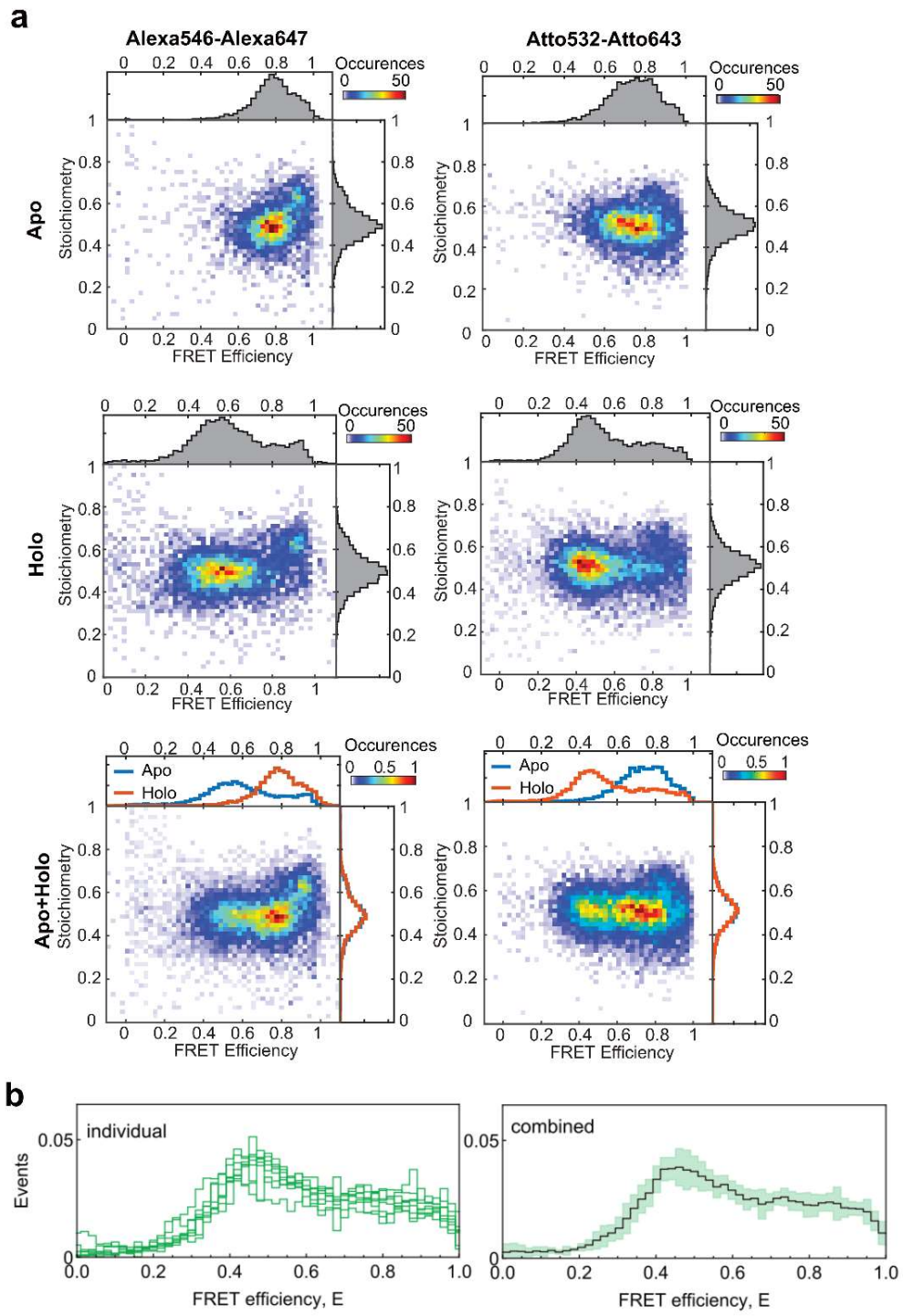
Supplementary Figure 13. Filtered fluorescence correlation spectroscopy curves for U2AF65. fFCS curves are shown for three different dye combinations, namely (a) Alexa546-Alexa647 (b)Atto532-Atto643 (c) Alexa488-Alexa647. In the analysis, a global fit of the two species autocorrelation functions, sACF, and two species cross-correlation functions, sCCF was performed. For simplicity, only one sACF function is shown. The fit model consisted of a diffusion term and two kinetic terms with corresponding relaxation times, $t_{R,1}$ and

$t_{R,2}$ (Supplementary Note 17). The obtained relaxation times (denoted as vertical lines) are consistent between different dye combinations as well as between apo and holo states.



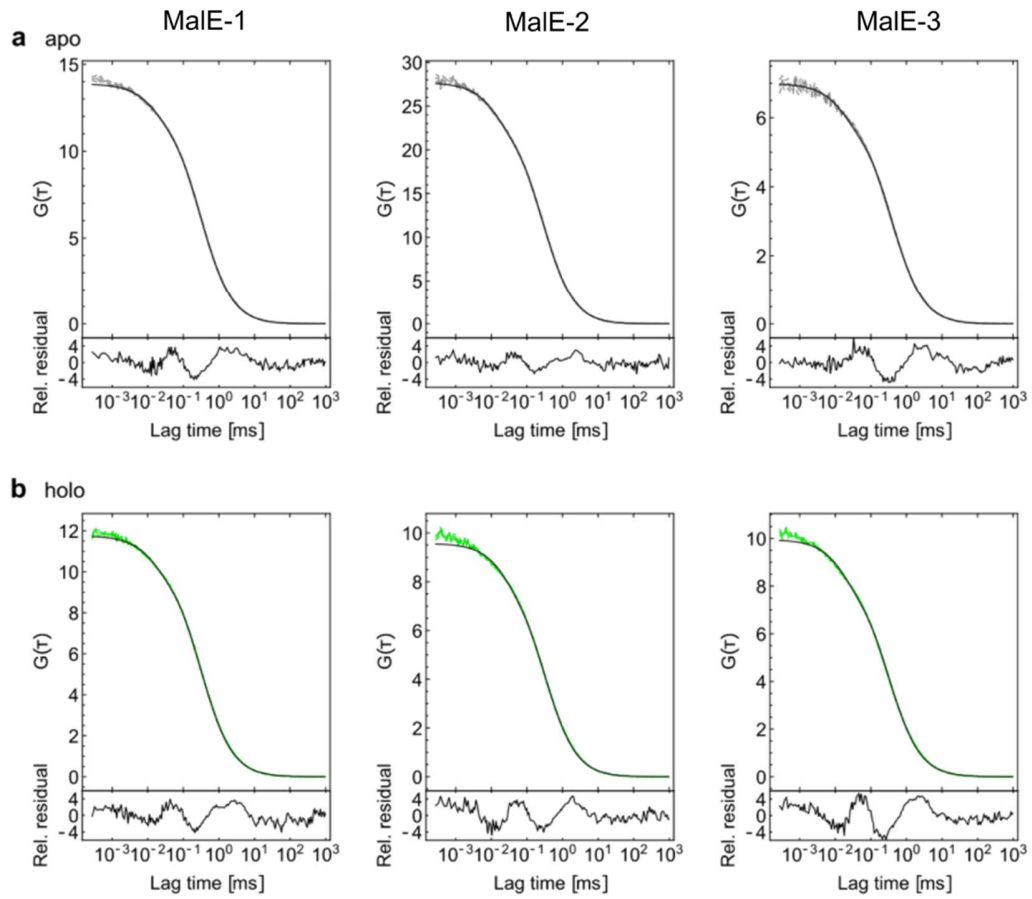
Supplementary Figure 14. A global dynamic photon distribution analysis (PDA) of apo and holo measurements of U2AF65 labeled with the Atto532-Atto643 dye pair. Different binning times (from top to bottom: 0.5 ms, 1 ms, 1.5 ms and 2.0 ms) for *left*: apo measurements (in grey) and *right*: holo measurement (in green) are shown. The analysis was performed globally over the apo and holo measurements using integration windows of 0.5, 1.0, 1.5 and 2.0 ms and shared distances for the compacted apo conformation (red), the detached apo ensemble (purple) and the holo state (blue). The dynamic interconversion between the compacted apo state and apo conformational ensemble is shown in yellow. An additional low-FRET population (magenta) had to be included, which most likely originates from photobleaching. All distances and kinetic rates for the apo ensemble are globally optimized

while the amplitudes of the apo and holo populations were kept constant within the integration time windows for the apo and holo states. See Supplementary Note 18 for details.



Supplementary Figure 15: The very-high-FRET population in U2AF65. **(a)** FRET efficiency versus stoichiometry plots are shown for two different dye combinations (Alexa546-Alexa647, left panel and Atto532-Atto643, right panel). These are plotted for apo (upper graph), holo (middle graph) and combined both apo and holo (lower graphs). A very-high-FRET populations with a different stoichiometry is visible in all measurements but with different amplitudes. **(b)** The FRET efficiency histograms (left) from the individual laboratories and (right) the combined histogram showing the mean (solid line) and a standard deviation (pale) after avoiding the subpopulation

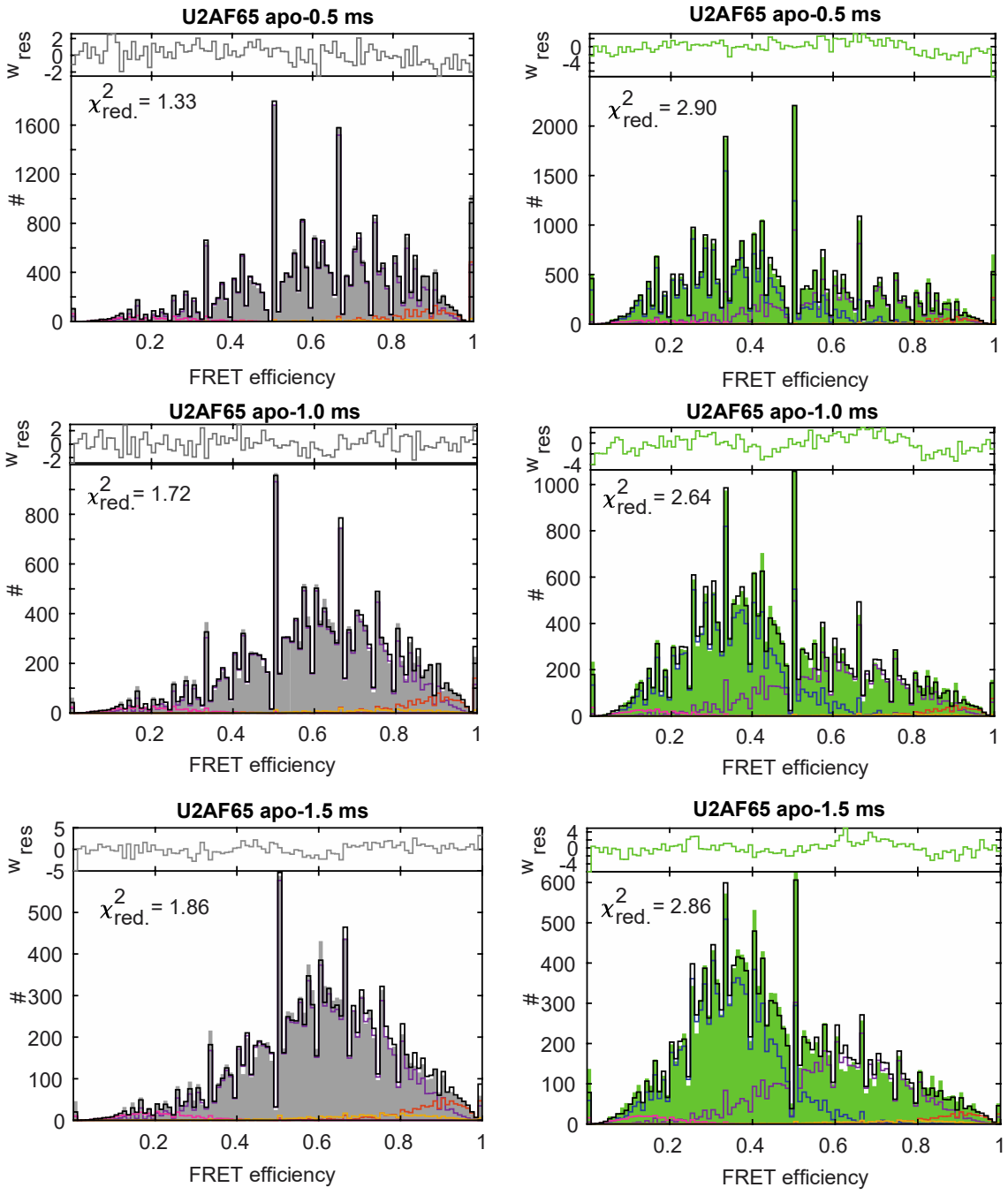
with slight acceptor quenching by using only bursts with a stoichiometry between 0.2-0.5 to build the smFRET histograms measured by the different labs.



Supplementary Figure 16. FCS experiments to exclude the presence of large aggregates. Fluorescence Correlation Spectroscopy (FCS) measurements for the mutants mutants MalE-1 (left), MalE-2 (middle), and MalE-3 (right) in **(a)** the apo state and **(b)** in the holo state are plotted. The two orthogonally oriented polarizations in the donor detection channel were correlated to remove the detector afterpulsing at short lag times (green line). All correlation curves were fitted with a standard model including a triplet fraction (black line):

$$G(t) = \frac{\gamma}{\langle N \rangle}\left(1+\frac{t}{\tau_{\text{diff}}}\right)^{-1}\left(1+\frac{t}{\tau_{\text{diff}}p^2}\right)^{-\frac{1}{2}}\left(1+\frac{T}{1-T}e^{-\frac{t}{\tau_{\text{trip}}}}\right),$$

where T is the triplet fraction, τ_{trip} the triplet lifetime, p is the structural or elongation factor of the confocal volume given as the ratio of the axial and lateral dimensions, τ_{diff} is the diffusion time, and $\gamma = 2^{-3/2}$ is the geometric factor[80]. The confocal instrument was calibrated using free Alexa546 dye with a published[81] diffusion coefficient of 341 $\mu m^2 s^{-1}$ for which a diffusion time of 95±10 µs was found. The overall diffusion time for the six measurements was 325±40 µs. None of the correlation curves show any indication of the presence of protein aggregates.



Supplementary Figure 17. A global dynamic photon distribution analysis (PDA) of the apo and holo measurements of U2AF65 labeled with the Alexa546-Alexa647 dye-pair. Left: apo measurement (in grey), right: holo measurement (in green). The analysis was performed globally over the apo and holo measurements for integration windows of (*top*) 0.5 ms, (*middle*) 1.0 ms, and (*bottom*) 1.5 ms using shared distances for the compacted apo conformation (red), the detached apo ensemble (purple) and the holo state (blue). The dynamic interconversion between the compacted apo state and the apo conformational ensemble is shown in yellow. All distances and kinetic rates for the apo ensemble are globally optimized while the amplitudes of the apo and holo populations were kept constant within the integration time windows for the apo and holo states. For details, see Supplementary Note 18.

Supplementary Tables

Supplementary Table 1. FRET efficiency correction factors reported by the participating labs for MalE constructs labeled with Alexa546-Alexa647

Overview of all correction factors for MalE and the resulting change in the FRET efficiency upon application of the correction factors is shown as an example for the holo state of the MalE-2 mutant. The correction factors: α is spectral crosstalk of donor fluorescence into the acceptor channel, ß is the normalization of direct donor and acceptor excitation fluxes, γ is the differences in donor and acceptor quantum yields and detection efficiencies and δ is the ratio of indirect and direct A excitation. $\langle E_{app} \rangle$ is the uncorrected FRET efficiency and $\langle E \rangle$ is the corrected FRET efficiency.

Lab#	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\delta}$	$\langle E_{app} \rangle$	$\langle E \rangle$
1	0.03	2.38	0.23	0.08	0.43	0.72
2	0.05	0.5	0.34	0.32	0.54	0.73
3	0.05	1.70	0.51	0.07	0.56	0.66
4	0.04	2.65	0.31	0.06	0.42	0.66
5	0.04	1.56	0.34	0.14	0.54	0.72
6	0.08	1.59	0.54	0.12	0.57	0.66
7	0.07	0.64	0.63	0.12	0.59	0.70
8	0.04	1.53	0.48	0.11	0.47	0.63
9	0.06	1.84	0.47	0.10	0.57	0.69
10	0.03	1.88	0.31	0.11	0.46	0.72
11	0.03	2.37	0.32	0.07	0.43	0.64
12	0.04	1.99	0.25	0.07	0.47	0.74
13	0.06	0.60	0.34	0.32	0.58	0.76
14	0.04	2.00	0.23	0.09	0.41	0.77
15	0.05	1.42	0.33	0.05	0.57	0.79
16	0.05	1.31	0.46	0.14	0.60	0.72
17 *	0.06	1.26	0.55	0.06		
18 **	0.01	4.86	0.09	0.10	0.18	0.59
19 ***	0.18				0.59	

*Due to measurement problems with a large bleaching contribution, data could not be fitted with a Gaussian distribution (the FRET population was small with most bursts containing photobleaching events). Data were not considered in the evaluation of the mean and standard deviation.

****T**he measurements were performed in a regime with a $\gamma < 0.1$, where the error in γ is significantly increased. Data were not considered in the evaluation of the mean and standard deviation.

***Due to measurement problems, data could not be corrected for direct excitation (δ) and γ. Data were not considered in the evaluation of mean and standard deviation.

Supplementary Table 2. FRET efficiency correction factors reported by the 8 labs for apo and holo measurements on U2AF65 labeled with Atto532-Atto643. α is the spectral crosstalk of donor fluorescence into the acceptor channel, ß is the normalization of direct donor and acceptor excitation fluxes, γ is the differences in donor and acceptor quantum yields and detection efficiencies and δ is the ratio of indirect and direct acceptor excitation.

Lab#	α	β	γ	δ
1	0.02	0.78	0.59	0.06
2	0.06	-	1.1	0.23
3	0.04	0.64	0.80	0.05
4	0.03	1.05	0.83	0.09
7	0.04	0.62	1.24	0.06
8	0.03	0.91	0.73	0.05
11	0.05	-	0.64	0.02
14	0.05	-	0.64	0.09

Supplementary Table 3. Reported mean FRET Efficiency and standard deviation for MalE samples by the participating laboratories: The mean FRET efficiency values $\langle E \rangle$ and distribution widths σ_E provided by the participating labs for all the three studied mutants of MalE labeled with Alexa546 and Alexa647 under both apo and holo conditions are listed. The difference in the reported mean FRET efficiency for individual laboratories between the apo and holo state was calculated as $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$. The calculated mean and standard deviation of the reported FRET efficiencies over all labs are given by $\mu_{\langle E \rangle}$ and $\sigma_{\langle E \rangle}$ and for the FRET efficiency difference, $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$, by $\boldsymbol{\mu}_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$ and $\sigma_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$, respectively.

	MalE-1					MalE-2					MalE-3				
Lab#	**Apo**		**Holo**		$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$	**Apo**		**Holo**		$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$	**Apo**		**Holo**		$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$
	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E		$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E		$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	
1	0.500	0.128	0.674	0.098	0.174	0.832	0.061	0.724	0.089	-0.108	0.908	0.039	0.922	0.035	0.014
2	0.520		0.700		0.180	0.840		0.730		-0.11	0.910		0.910		0
3	0.453	0.096	0.641	0.077	0.188	0.794	0.058	0.654	0.072	-0.14	0.904	0.040	0.915	0.039	0.011
4	0.461	0.169	0.637	0.119	0.176	0.785	0.096	0.658	0.126	-0.127	0.896	0.066	0.901	0.064	0.005
5	0.522		0.703		0.181	0.845		0.724		-0.121	0.925		0.912		-0.013
6	0.509	0.129	0.644	0.109	0.135	0.820	0.073	0.657	0.090	-0.163	0.900	0.054	0.911	0.060	0.011
7	0.454	0.119	0.641	0.086	0.187	0.807	0.058	0.702	0.068	-0.105	0.917	0.046	0.914	0.044	-0.003
8	0.414	0.130	0.602	0.105	0.188	0.771	0.057	0.627	0.101	-0.144	0.881	0.044	0.890	0.041	0.009
9	0.451	0.147	0.622	0.118	0.171	0.820	0.069	0.685	0.095	-0.135	0.911	0.054	0.911	0.053	0
10	0.444	0.180	0.631	0.144	0.187	0.834	0.081	0.713	0.102	-0.121	0.923	0.062	0.937	0.073	0.014
11	0.4	0.112	0.610	0.087	0.210	0.781	0.063	0.644	0.075	-0.137	0.876	0.054	0.887	0.047	0.011
12	0.523	0.118	0.706	0.081	0.183	0.849	0.052	0.74	0.073	-0.109	0.921	0.039	0.924	0.037	0.003
13	0.547	0.151	0.723	0.105	0.176	0.872	0.065	0.761	0.093	-0.111	0.942	0.051	0.952	0.043	0.010
14	0.580		0.745		0.165	0.865		0.765		-0.1	0.935		0.970		0.035
15	0.621	0.114	0.76	0.072	0.139	0.877	0.047	0.787	0.066	-0.09	0.936	0.035	0.936	0.034	0
16	0.466	0.100	0.655	0.082	0.189	0.831	0.050	0.716	0.067	-0.115	0923	0.038	0.93	0.038	0.007
$\boldsymbol{\mu}_{\langle E \rangle}$ or $\boldsymbol{\mu}_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$	0.492		0.667		0.177	0.826		0.705		-0.121	0.913		0.920		0.007
$\boldsymbol{\sigma}_{\langle E \rangle}$ or $\boldsymbol{\sigma}_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$	0.060		0.049		0.019	0.032		0.047		0.019	0.019		0.021		0.010

Supplementary Table 4. Reported and reanalyzed mean FRET efficiency and standard deviations for the U2AF65 experiments:

The U2AF65 protein was labeled with Atto532-Atto643 and labs provided the number of events for FRET efficiency over 51 bins between 0 and 1. To obtain the mean FRET efficiency $\langle E \rangle$ and distribution width σ_E values for U2AF65, the reported events were fitted with one and two Gaussian distribution functions for apo and holo conditions respectively. For reanalysis, the original datasets from the individual labs were obtained and reanalyzed according to the procedure detailed in the Supplementary Note 4. U2AF65 under holo conditions has a major fraction of holo state and a minor fraction of apo state due to incomplete saturation of the RNA ligand. Only 8 labs participated

Lab #	Reported						Reanalyzed					
	Apo		+RNA				Apo		+RNA			
			Holo state		Apo state				Holo state		Apo state	
	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E
1	0.748	0.192	0.436	0.150	0.756	0.182	0.736	0.200	0.405	0.158	0.769	0.243
2	0.749	0.202	0.455	0.168	0.811	0.174	0.757	0.207	0.432	0.157	0.800	0.208
3	0.708	0.228	0.506	0.206	0.785	0.142	0.738	0.262	0.449	0.163	0.767	0.229
4	0.728	0.205	0.414	0.197	0.773	0.152	0.736	0.211	0.385	0.158	0.768	0.179
7	0.857	0.205	0.495	0.196	0.875	0.181	0.800	0.280	0.495	0.196	0.875	0.181
8	0.795	0.169	0.521	0.206	0.857	0.117	0.742	0.217	0.414	0.192	0.814	0.216
11	0.729	0.204	0.423	0.133	0.727	0.231	0.736	0.196	0.429	0.112	0.707	0.269
14	0.718	0.199	0.430	0.164	0.743	0.195	0.749	0.214	0.448	0.131	0.755	0.241
$\mu_{\langle E \rangle}$	0.754		0.460		0.790		0.749		0.432		0.782	
$\sigma_{\langle E \rangle}$	0.049		0.042		0.053		0.023		0.033		0.049	

in the U2AF65 study because of its complexity. The calculated mean and standard deviation of the reported and reanalyzed FRET efficiencies over all labs for each sample are given by $\mu_{\langle E \rangle}$ and $\sigma_{\langle E \rangle}$ respectively.

Supplementary Table 5. The statement of the individual laboratories regarding the dynamics of MalE and U2AF65.

MalE-1: 29C/352C, MalE-2: 87C/186C; MalE-3: 134C/186C. "-" = no statement, "n/a"=not applicable due to experimental limitations (instrumentation & established evaluation routines).

Lab#	Method	Sub-ms dynamics in MalE-1/2/3	Sub-ms dynamics in U2AF65 (apo / holo)
1	BVA + E-Tau	no / no / no	yes / yes
2	E-Tau (fFCS, PDA)	yes / yes / yes	yes / yes
3	BVA + E-Tau	no / no / no	yes / yes
4	BVA + E-Tau	no / no / no	yes / yes
5	n/a	n/a	n/a
6	E-Tau (fFCS)	no / no / -	yes / yes
7	BVA	no / - / yes	yes / yes
8	BVA	no / no / no	yes / yes
9	BVA + E-Tau	no / no / no	n/a
10	BVA	no / no / no	n/a

11	BVA + E-Tau	no / no / no	yes / yes
12	BVA	no / no / no	n/a
13	fFCS	no / no / yes	n/a
14	BVA + E-Tau	no / no / no	n/a
15	n/a	n/a	n/a
16	BVA	no / no / no	n/a
17	n/a	n/a	n/a
18	n/a	n/a	n/a
19	n/a	n/a	n/a

Supplementary Table 6. The apparent dynamic shifts determined for both MalE and U2AF65 samples for the data collected from 8 labs. MalE and U2AF65 samples were labeled with Alexa546-Alexa647 and Atto532-Atto643 dye pairs respectively. The apparent dynamic shifts for the data collected in different labs are shown below for both BVA and E-τ plots. The dynamic shift, *ds* of the peak of the population was determined graphically as explained in Supplementary Note 6. For U2AF65 in the holo state, the dynamic shift was assessed only for the low-FRET RNA-bound population. Note the negative dynamic shifts which occur due to dye artifacts.

Lab#			1	2	3	4	7	8	11	14
BVA	MalE-1	Apo	0.0049	0.0007	0.0128	-0.0027	0.0167	0.0075	0.0042	0.0093
		Holo	0.0083	0.0022	0.0135	-0.001	0.0166	0.0073	0.0079	0.0098
	MalE-2	Apo	0.0083	0.0009	0.014	0.0023	0.0167	0.0103	0.0080	0.0105
		Holo	0.0070	0.0025	0.0143	0.0004	0.021	0.008	0.0062	0.0097
	MalE-3	Apo	0.0127	-0.0002	0.0081	-0.0043	0.0166	0.01	0.0107	0.0293
		Holo	0.0068	-0.0036	0.0087		-0.0043	0.0024	0.0077	0.0028
	MalE-4	Apo	0.0022		-0.0031					
		Holo	0.004		0.0070					
	MalE-5	Apo	0.0016		0.0031					
		Holo	0.003		0.0009					
	U2AF65	Apo	0.0283	0.0206	0.0217	0.026		0.0261	0.0363	0.0256
		Holo	0.0131	0.018	0.0259	0.0117		0.0178	0.0160	0.0171
E-τ	MalE-1	Apo	0.074	0.059	0.055	0.003		0.034		0.088
		Holo	0.069	0.057	0.062	0.004		0.031		0.068
	MalE-2	Apo	0.003	-0.007	0.018	0.01		0.029		0.042
		Holo	0.027	-0.002	0.025	0.015		0.011		0.05
	MalE-3	Apo	0.031	-0.013	0.026	0.004		0.044		0.043
		Holo	0.019	-0.019	0.037	0.035		0.027		0.046
	MalE-4	Apo	0.025	0.014	0.02					
		Holo	0.055	0.03	0.042					
	MalE-5	Apo	0.042	0.027	0.03					
		Holo	0.036	0.018	0.026					
	U2AF65	Apo	0.167	0.134	0.097	0.122		0.141		0.159
		Holo	0.022	0.008	0.019	0.035		0.012		0.029

Supplementary Table 7. The apparent dynamic shift values for different dye combinations of MalE and U2AF65 FRET variants as determined by three labs.

The estimated distance fluctuation δR and FRET-averaged distance $R_{\langle E \rangle}$ is given for the measurements that passed the filtering procedure based on the estimated distance uncertainty from the residual anisotropies. Furthermore, the *ds* value for DNA rulers for each construct (LF: low-FRET, MF: medium-FRET, HF: high-FRET) and dye labels are provided and well as the average value over all measurement with the given error being the standard error of the mean, SEM.

			Lab#1			*Lab#2*			*Lab#3*		
Sample	**Dye combination**	**state**	***ds***	$R_{\langle E \rangle}$ **[Å]**	δR **[Å]**	***ds***	$R_{\langle E \rangle}$ **[Å]**	δR **[Å]**	***ds***	$R_{\langle E \rangle}$ **[Å]**	δR **[Å]**
MalE-1	Alexa546-Alexa647	apo	0.074			0.059			0.055		
MalE-1	Alexa546-Abb. STAR635P	apo	0.016			0.033					
MalE-1	Atto532-Atto643	apo	0.021	65.5	6.6						
MalE-1	Alexa546-Alexa647	holo	0.069			0.057			0.062		
MalE-1	Alexa546-Abb. STAR635P	holo	0.003			0.035					
MalE-1	Atto532-Atto643	holo	0.020	57.6	4.5						
MalE-2	Alexa546-Alexa647	apo	0.003	49.8	1.6	-0.007	49.3	0.2	0.018	51.2	3.9
MalE-2	Atto532-Atto643	apo	0.007	47.3	2.2	0.009	47.3	2.5			
MalE-2	Alexa488-Alexa647	apo	0.001	47.9	0.8	-0.002	47.9	0.2			
MalE-2	Alexa546-Alexa647	holo	0.027	55.1	4.8	-0.002	55.3	0.0	0.025	56.3	4.7
MalE-2	Atto532-Atto643	holo	0.020	54.4	4.0	0.018	54.4	3.8			
MalE-2	Alexa488-Alexa647	holo	0.010	54.1	3.3	0.009	54.1	3.1			
MalE-3	Alexa546-Alexa647	apo	0.031	44.2	5.7	-0.013	44.4	0.0	0.026	43.6	5.3
MalE-3	Atto532-Atto643	apo	0.003	39.3	1.7	0.002	39.3	1.4			
MalE-3	Alexa488-Alexa647	apo	-0.037	40.4	0.0	0.013	40.4	2.6			
MalE-3	Alexa546-Alexa647	holo	0.019	44.2	4.5	-0.019	43.1	0.0	0.037	43.8	6.3
MalE-3	Atto532-Atto643	holo	-0.009	40.1	0.0	-0.007	40.1	0.0			
MalE-3	Alexa488-Alexa647	holo	-0.013	40.4	0.0	0.002	40.4	1.1			
MalE-4	Alexa546-Alexa647	apo	0.025			0.014			0.02		
MalE-4	Atto532-Atto643	apo	-0.005	66.9	0.0				0.025	66.9	7.8
MalE-4	Alexa488-Alexa647	apo	0.001	66.2	2.0	0.024	66.2	10.8			
MalE-4	Alexa546-Alexa647	holo	0.055			0.03			0.042		
MalE-4	Atto532-Atto643	holo	0.007	57.7	2.7				0.045	57.7	6.8
MalE-4	Alexa488-Alexa647	holo	-0.001	54.9	0.0	0.026	54.9	5.5			
MalE-5	Alexa546-Alexa647	apo	0.042			0.027			0.03		
MalE-5	Atto532-Atto643	apo	0.014	60.6	4.2	0.038	59.4	6.6	0.031	59.4	6.0
MalE-5	Alexa488-Alexa647	apo	0.032	60.4	8.4	0.022	60.4	7.1			
MalE-5	Alexa546-Alexa647	holo	0.036			0.018			0.026		
MalE-5	Atto532-Atto643	holo	0.032	49.1	4.7	0.027	49.2	4.3	0.04	49.2	5.3
MalE-5	Alexa488-Alexa647	holo	0.017	49.3	3.4	0.033	49.3	4.8			
U2AF	Alexa546-Alexa647	apo	0.166			0.129					
U2AF	Atto532-Atto643	apo	0.167			0.164					
U2AF	Alexa488-Alexa647	apo	0.168			0.128					
U2AF	Alexa546-Alexa647	holo	0.033			0.024					
U2AF	Atto532-Atto643	holo	0.022			0.025					
U2AF	Alexa488-Alexa647	holo	0.053			0.012					
DNA	Alexa488-Atto647N	LF				0.007	75.0	13.9			
DNA	Alexa488-Cy5	LF				0.015	73.6	13.9			
DNA	Alexa488-Alexa647	LF				0.012	72.5	11.9			

DNA	Alexa488-Atto647N	MF				-0.008	61.5	0.3			
DNA	Alexa488-Cy5	MF				0.011	61.2	5.4			
DNA	Alexa488-Alexa647	MF				-0.003	62.4	0.0			
DNA	Alexa488-Atto647N	HF				-0.012	51.7	0.0			
DNA	Alexa488-Cy5	HF				0.017	51.5	3.7			
DNA	Alexa488-Alexa647	HF				-0.016	52.4	0.0			
DNA	average over different dye combinations	-				0.0026 ± 0.0044					

Supplementary Table 8. The expected FRET efficiencies and dynamic shifts for the different experimental systems based on structural models from the PDB for the dye pair Alexa546-Alexa647. FRET efficiencies were predicted by AV simulations using the parameters given in Supplementary Table 19 based on the following PDB IDs: MalE apo - 1OMP, MalE holo - 1ANF, U2AF65 apo - 2YHO, U2AF65 holo - 2YH1. The expected dynamic shifts were calculated as described in Supplementary Note 7 using Eq. 7.1 for the E-τ plot and Eq. 7.23 for BVA.

	FRET efficiency, E		Expected dynamic shift, ds	
Sample	**apo**	**holo**	**E-τ plot**	**BVA**
MalE-1	0.358	0.582	0.0169	0.0215
MalE-2	0.827	0.695	0.0131	0.0089
MalE-3	0.947	0.941	0.0002	0.0000
MalE-4	0.443	0.680	0.0231	0.0240
MalE-5	0.578	0.826	0.0382	0.0282
U2AF65	0.882	0.372	0.1426	0.0984

Supplementary Table 9. Global fit of polarization-resolved fluorescence decays from ensemble measurements of single mutants of MalE.

Rotational correlation times, ρ_j, with corresponding amplitudes, b_j, for Alexa647 and Alexa546 dyes from ensemble TCSPC measurements of single cysteine MalE mutants are reported (Supplementary Note 9, Eqn. 9.8). The rotational correlation times and fluorescence lifetimes as well their corresponding amplitudes were obtained from a global fit of the polarization-resolved fluorescence decays, $f_{VV}(t)$ and $f_{VH}(t)$, as described in Supplementary Note 9. The fit quality was judged by χ_r^2. The fit results indicate strong sticking interactions for positions K29C, A134C and S352C for the Alexa546 dye while, for the Alexa647 fluorophore, pronounced sticking interactions were found only at position S352C. From the fluorescence lifetime analysis, it can be seen that positions D87C and S352C are prone to quenching for the donor dye. Fitted polarization-resolved decays are combined into anisotropy decay and displayed in Fig. 5b for two representative mutation sites, namely S352C and K29C. Furthermore, obtained residual anisotropies are used in computation of Accessible Contact Volumes (Fig. 5c), which improved the agreement between modelled and measured distances (Fig. 5d)

Dye	Mutant	Condition	ρ_1 [ns]	$b_1 = r_{\infty,tr}$	ρ_2 [ns]	b_2	ρ_3 [ns]	b_3	τ_1 [ns]	x_1	τ_2 [ns]	x_2	χ_r^2
Alexa 546	K29C	apo	26	0.332	1.193	0.048			4.03	1.00			1.01
	D87C	apo	20	0.206	3.500	0.087	0.572	0.087	3.85	1.00			1.03
	A134C	apo	20	0.342			0.138	0.038	3.93	1.00			0.99
	A186C	apo	20	0.146	3.187	0.089	0.580	0.145	3.94	1.00			1.07
	S352C	apo	20	0.272	2.442	0.053	0.594	0.055	3.98	0.90	1.8	0.10	1.05
	K29C	holo	20	0.316			0.413	0.064	4.05	1.00			0.90
	D87C	holo	20	0.236	1.198	0.144			3.91	1.00			0.95
	A134C	holo	20	0.343			0.477	0.037	3.99	1.00			0.94
	A186C	holo	20	0.161			0.973	0.218	3.93	1.00			0.94
	S352C	holo	20	0.276	1.100	0.104			4.02	0.90	1.94	0.10	0.92
Alexa 647	K29C	apo	>50	0.125			0.639	0.255	1.878	0.14	1.165	0.86	1.19
	D87C	apo	20	0.169			0.748	0.210	1.729	0.18	1.174	0.82	1.07
	A134C	apo	20	0.170			0.539	0.210	1.833	0.18	1.205	0.82	1.15
	A186C	apo	20	0.082	1.108	0.189	0.428	0.109	1.557	0.17	1.145	0.83	1.05
	S352C	apo	30	0.258			0.695	0.122	1.733	0.42	1.203	0.58	1.10
	K29C	holo	20	0.138			0.578	0.241	1.885	0.13	1.169	0.87	1.22
	D87C	holo	20	0.167			0.691	0.213	1.745	0.18	1.174	0.82	1.11
	A134C	holo	20	0.164			0.536	0.216	1.817	0.18	1.202	0.82	1.17
	A186C	holo	20	0.084			0.756	0.296	1.660	0.09	1.172	0.91	1.03
	S352C	holo	20	0.283			0.498	0.097	1.757	0.42	1.209	0.58	1.18

Supplementary Table 10. Steady-state and residual time-resolved anisotropy values of single mutant MalE samples labelled with Alexa 546 and Alexa 647, respectively. In addition to time-resolved anisotropies (Supplementary Table 9), for comparison we also give the steady-state anisotropy of the same fluorophores as free dyes and coupled to double-stranded DNA. Residual anisotropies obtained from time-resolved analysis (Supplementary Note 9, Eqn. 9.4 - 9.11) of donor and acceptor dye are used in computation of Accessible Contact Volumes (ACVs) (Fig. 5c), which improved the agreement between modelled and measured distances (Fig. 5d)

	$D_{ex}\|D_{em}$ channels (donor dye)		$A_{ex}\|A_{em}$ channels (acceptor dye)	
Sample	**Steady-state anisotropy, r_{ss}**	**Residual anisotropy, $r_{\infty,tr}$**	**Steady-state anisotropy, r_{ss}**	**Residual anisotropy, $r_{\infty,tr}$**
Free dye				
Alexa 546	0.035±0.003	0.01±0.02	-	
Alexa647	-		0.120±0.007	0.02±0.02
DNA-Standards				
8 base-pairs	0.114±0.003		0.184±0.012	
33 base-pairs	0.134±0.002		0.159±0.011	
Donor-only strand	0.134±0.003		-	
Acceptor-only strand	-		0.172±0.010	
Protein single mutants				
K29C, apo	0.285±0.017	0.332	0.198±0.015	0.125
K29C, holo	0.280±0.017	0.316	0.199±0.018	0.138
D87C, apo	0.231±0.012	0.206	0.217±0.016	0.169
D87C, holo	0.225±0.005	0.236	0.229±0.017	0.167
A134C, apo	0.290±0.016	0.342	0.215±0.019	0.170
A134C, holo	0.281±0.007	0.343	0.216±0.003	0.164
A186C, apo	0.176±0.018	0.146	0.186±0.014	0.082
A186C, holo	0.161±0.010	0.161	0.186±0.016	0.084
S352C, apo	0.247±0.007	0.272	0.272±0.015	0.258
S352C holo	0.243±0.010	0.276	0.263±0.002	0.283

Supplementary Table 11. Global fit of the polarization-resolved and magic-angle fluorescence decays from sub-ensemble data of MalE and U2AF65 samples. The different rotational correlation times, ρ_j, with corresponding amplitudes, b_j, for the $A_{ex}|A_{em}$ channels (acceptor dye) and $D_{ex}|D_{em}$ channels (donor dye) on a sub-ensemble DA population describe different depolarization processes. The fluorescence lifetimes, τ_i, and fraction of molecules with a given lifetime, x_i, were obtained from the magic-angle decay function, which was fitted globally with the polarization-resolved fluorescence decays, $f_{VV}(t)$ and $f_{VH}(t)$, as described in Supplementary Note 9, Eqn. 9.4 - 9.11. The fit quality was judged by the reduced Chi-squared value, χ_r^2. Fitted residual anisotropies for different dye combinations are used in the calculation of relative distance uncertainties (Fig. 5e), and it is shown that they are correlated to the observed dynamic shift (Fig. 5f).

	$A_{ex}\|A_{em}$ channels										
Sample	**Dye combination**	**Condition**	ρ_1 [ns]	$b_1 = r_{\infty,tr}$	ρ_2 [ns]	b_2	τ_1 [ns]	x_1	τ_2 [ns]	x_2	χ_r^2
MalE-1	Alexa546-Alexa647	apo	>50	0.331	0.84	0.049	1.69	0.50	0.93	0.50	0.96
	Alexa546-AbberiorSTAR635P	apo	20	0.209	0.16	0.171	4.00	1.00			0.98
	Atto532-Atto643	apo	20	0.181	0.27	0.199	3.95	1.00			0.99
	Alexa546-Alexa647	holo	20	0.314	0.10	0.066	1.66	0.57	0.95	0.43	1.39
	Alexa546-Abberior STAR635P	holo	20	0.182	0.12	0.198	4.00	1.00			0.98
	Atto532-Atto643	holo	20	0.212	0.32	0.168	3.80	1.00			1.15
MalE-2	Alexa546-Alexa647	apo	20	0.212	0.24	0.168	1.39	0.61	0.75	0.39	1.15
	Atto532-Atto643	apo	20	0.160	0.33	0.220	3.60	1.00			1.08
	Alexa488-Alexa647	apo	>50	0.150	0.57	0.230	1.30	1.00			0.94
	Alexa546-Alexa647	holo	20	0.228	0.25	0.152	1.48	0.48	0.91	0.52	1.10
	Atto532-Atto643	holo	20	0.186	0.45	0.194	6.30	1.00			1.00
	Alexa488-Alexa647	holo	20	0.150	0.49	0.230	1.72	0.21	1.22	0.79	1.11
MalE-3	Alexa546-Alexa647	apo	20	0.215	0.27	0.165	1.37	0.58	0.75	0.42	1.09
	Atto532-Atto643	apo	20	0.149	0.31	0.231	3.78	1.00			1.1
	Alexa488-Alexa647	apo	20	0.192	0.62	0.188	1.37	1.00			1.08
	Alexa546-Alexa647	holo	20	0.201	0.28	0.179	1.45	0.49	0.82	0.51	1.16
	Atto532-Atto643	holo	20	0.147	0.28	0.233	3.67	1.00			1.12
	Alexa488-Alexa647	holo	20	0.185	0.44	0.195	1.33	1.00			1.16
MalE-4	Alexa546-Alexa647	apo	>50	0.205	0.59	0.175	1.30	0.68	0.60	0.32	0.83
	Atto532-Atto643	apo	20	0.144	0.42	0.236	3.84	1.00			1.09
	Alexa488-Alexa647	apo	20	0.113	0.29	0.267	1.50	0.58	0.88	0.42	1.16
	Alexa546-Alexa647	holo	20	0.217	0.32	0.163	1.67	0.27	0.91	0.73	0.9
	Atto532-Atto643	holo	20	0.140	0.29	0.240	3.53	0.84	5.08	0.16	1.07
	Alexa488-Alexa647	holo	20	0.121	0.30	0.259	1.84	0.23	1.12	0.77	1.16
MalE-5	Alexa546-Alexa647	apo	>50	0.275	0.68	0.105	1.65	0.24	1.00	0.76	1.20
	Atto532-Atto643	apo	20	0.198	0.24	0.182	3.70	1.00			1.04
	Alexa488-Alexa647	apo	20	0.160	0.24	0.220	1.54	0.43	1.06	0.57	1.16
	Alexa546-Alexa647	holo	>50	0.296	1.20	0.084	1.87	0.16	1.00	0.84	1.08
	Atto532-Atto643	holo	20	0.201	0.41	0.179	3.40	1.00			1.16
	Alexa488-Alexa647	holo	20	0.108	0.34	0.272	1.68	0.38	1.00	0.62	1.11
U2AF65	Alexa546-Alexa647	apo	21	0.292	0.10	0.088	1.70	0.82	0.87	0.18	0.98
	Atto532-Atto643	apo	4	0.128	0.03	0.252	4.27	1.00			0.99
	Alexa488-Alexa647	apo	18	0.208	0.14	0.172	1.49	0.92	3.40	0.08	1.03

	Alexa546-Alexa647	holo	24	0.230	0.10	0.150	1.70	0.72	1.00	0.28	1.05
	Atto532-Atto643	holo	3.4	0.117	0.06	0.263	4.10	1.00			1.00
	Alexa488-Alexa647	holo	18	0.204	0.20	0.176	1.41	0.92	2.90	0.08	1.05

$D_{ex}\|D_{em}$ channels															
Sample	**Dye pair**	**Condition**	ρ_1 [ns]	b_1= $r_{\infty,tr}$	ρ_2 [ns]	b_2	ρ_3 [ns]	b_3	τ_1 [ns]	x_1	τ_2 [ns]	x_2	τ_3 [ns]	x_3	χ_r^2
MalE-1	Alexa546-Alexa647	apo	80	0.340	0.63	0.040			2.73	0.52	1.12	0.48			1.16
	Alexa546-Abberior STAR635P	apo	>50	0.327	1.60	0.050			2.87	0.49	1.18	0.51			1.03
	Atto532-Atto643	apo	20	0.229	0.21	0.151			2.22	0.79	3.62	0.21			1.24
	Alexa546-Alexa647	holo	46	0.329	0.16	0.050			1.69	0.48	0.77	0.41	3.09	0.11	1.19
	Alexa546-Abberior STAR635P	holo	>50	0.341	1.61	0.039			2.38	0.38	0.78	0.62			1.19
	Atto532-Atto643	holo	20	0.221	0.28	0.159			2.25	0.64	0.91	0.36			1.19
MalE-2	Alexa546-Alexa647	apo	>50	0.252	4.96	0.128			0.95	0.42	0.30	0.53	3.06	0.05	1.2
	Atto532-Atto643	apo	20	0.198	0.26	0.182			0.58	0.67	1.70	0.33			1.1
	Alexa488-Alexa647	apo	20	0.136	0.19	0.244			0.95	0.54	2.30	0.46			1.1
	Alexa546-Alexa647	holo	20	0.255	0.37	0.125			1.30	0.51	0.38	0.44	3.26	0.05	1.01
	Atto532-Atto643	holo	20	0.207	0.15	0.173			2.02	0.50	0.77	0.50			1.10
	Alexa488-Alexa647	holo	20	0.133	0.23	0.247			0.70	0.27	2.5	0.73			1.14
MalE-3	Alexa546-Alexa647	apo	20	0.255	0.26	0.125			0.33	0.81	1.04	0.18	6.30	0.01	1.03
	Atto532-Atto643	apo	20	0.048	0.14	0.332			0.45	0.87	2.02	0.13			1.00
	Alexa488-Alexa647	apo	20	0.087	0.10	0.293			0.66	0.69	2.10	0.31			1.14
	Alexa546-Alexa647	holo	20	0.187	0.25	0.193			0.33	0.78	1.02	0.19	3.20	0.03	1.13
	Atto532-Atto643	holo	20	0.044	0.15	0.336			0.45	0.87	2.24	0.13			1.09
	Alexa488-Alexa647	holo	20	0.116	0.15	0.264			0.47	0.64	1.80	0.36			0.96
MalE-4	Alexa546-Alexa647	apo	>50	0.344	0.57	0.036			2.60	0.69	0.80	0.31			1.04
	Atto532-Atto643	apo	20	0.126	1.40	0.05	0.10	0.202	2.82	0.8	1.17	0.20			1.13
	Alexa488-Alexa647	apo	20	0.13	0.15	0.250			3.30	0.87	1.53	0.13			1.18
	Alexa546-Alexa647	holo	>50	0.348	1.35	0.032			2.26	0.38	0.80	0.62			1.00
	Atto532-Atto643	holo	20	0.166	0.10	0.214			1.64	0.57	0.40	0.23	3.00	0.2	1.12
	Alexa488-Alexa647	holo	20	0.116	0.13	0.264			2.80	0.68	1.09	0.32			1.02
MalE-5	Alexa546-Alexa647	apo	>50	0.335	0.81	0.04			2.35	0.49	0.80	0.50			1.07
	Atto532-Atto643	apo	20	0.238	0.15	0.142			2.36	0.58	0.60	0.42			1.11
	Alexa488-Alexa647	apo	20	0.143	0.13	0.236			3.00	0.85	1.11	0.14			1.14
	Alexa546-Alexa647	holo	>50	0.338	0.29	0.042			1.65	0.34	0.40	0.66			1.1
	Atto532-Atto643	holo	20	0.233	0.15	0.147			1.70	0.42	0.50	0.58			1.16
	Alexa488-Alexa647	holo	20	0.118	0.10	0.262			2.37	0.51	0.74	0.49			1.07
U2AF65	Alexa546-Alexa647	apo	22	0.276	0.10	0.104			2.50	0.39	0.60	0.61			1.17
	Atto532-Atto643	apo	18	0.253	0.21	0.127			2.59	0.41	0.53	0.59			1.27
	Alexa488-Alexa647	apo	18	0.200	0.19	0.180			2.95	0.44	0.67	0.56			1.16
	Alexa546-Alexa647	holo	25	0.294	0.10	0.086			2.49	0.43	0.77	0.57			1.14
	Atto532-Atto643	holo	18	0.25	0.17	0.130			2.48	0.57	0.65	0.43			1.29
	Alexa488-Alexa647	holo	18	0.18	0.17	0.200			2.84	0.67	0.73	0.33			1.09

Supplementary Table 12. Combined residual anisotropies of additional dye combinations for double-labeled MalE and U2AF65 samples. Residual anisotropies of different donor and acceptor pairs used for labeling of MalE and U2AF65 variants. Residual anisotropies are computed using two approaches: from time-resolved analysis of the polarization-resolved fluorescence decays and from steady-state anisotropy measurements using a two-component Perrin equation (Supplementary Note 9, Eqn. 9.1 - 9.3). The results from the two approaches were averaged. Furthermore, using the residual anisotropies of donor and acceptor fluorophores, a combined residual anisotropy for a given FRET pair is obtained (Supplementary Note 9, Eqn. 9.12 - 9.15).

Sample	Dye combination	state	$\langle r_{\infty,D}\rangle_{tr,ss}$	$\langle r_{\infty,A}\rangle_{tr,ss}$	$\langle r_{c,\infty}\rangle_{tr,ss}$
MalE-1	Alexa546-Alexa647	apo	0.350 ± 0.014	0.286 ± 0.064	0.316 ± 0.036
MalE-1	Alexa546-Abb. STAR635P	apo	0.336 ± 0.013	0.228 ± 0.027	0.277 ± 0.017
MalE-1	Atto532-Atto643	apo	0.194 ± 0.049	0.179 ± 0.004	0.186 ± 0.023
MalE-1	Alexa546-Alexa647	holo	0.339 ± 0.014	0.272 ± 0.059	0.304 ± 0.034
MalE-1	Alexa546-Abb. STAR635P	holo	0.336 ± 0.006	0.222 ± 0.056	0.273 ± 0.035
MalE-1	Atto532-Atto643	holo	0.200 ± 0.029	0.189 ± 0.033	0.195 ± 0.022
MalE-2	Alexa546-Alexa647	apo	0.237 ± 0.021	0.176 ± 0.050	0.204 ± 0.031
MalE-2	Atto532-Atto643	apo	0.154 ± 0.063	0.141 ± 0.027	0.147 ± 0.033
MalE-2	Alexa488-Alexa647	apo	0.110 ± 0.037	0.190 ± 0.056	0.144 ± 0.032
MalE-2	Alexa546-Alexa647	holo	0.257 ± 0.002	0.170 ± 0.082	0.209 ± 0.050
MalE-2	Atto532-Atto643	holo	0.161 ± 0.065	0.152 ± 0.048	0.156 ± 0.040
MalE-2	Alexa488-Alexa647	holo	0.117 ± 0.023	0.195 ± 0.063	0.151 ± 0.029
MalE-3	Alexa546-Alexa647	apo	0.257 ± 0.003	0.168 ± 0.066	0.208 ± 0.041
MalE-3	Atto532-Atto643	apo	0.035 ± 0.019	0.139 ± 0.014	0.069 ± 0.019
MalE-3	Alexa488-Alexa647	apo	0.090 ± 0.004	0.211 ± 0.026	0.137 ± 0.009
MalE-3	Alexa546-Alexa647	holo	0.212 ± 0.036	0.155 ± 0.065	0.181 ± 0.041
MalE-3	Atto532-Atto643	holo	0.048 ± 0.006	0.139 ± 0.012	0.082 ± 0.006
MalE-3	Alexa488-Alexa647	holo	0.104 ± 0.017	0.206 ± 0.030	0.147 ± 0.016
MalE-4	Alexa546-Alexa647	apo	0.345 ± 0.002	0.199 ± 0.009	0.262 ± 0.006
MalE-4	Atto532-Atto643	apo	0.140 ± 0.020	0.136 ± 0.011	0.138 ± 0.011
MalE-4	Alexa488-Alexa647	apo	0.116 ± 0.020	0.167 ± 0.077	0.139 ± 0.034
MalE-4	Alexa546-Alexa647	holo	0.352 ± 0.006	0.235 ± 0.025	0.288 ± 0.016
MalE-4	Atto532-Atto643	holo	0.165 ± 0.001	0.130 ± 0.015	0.146 ± 0.008
MalE-4	Alexa488-Alexa647	holo	0.110 ± 0.008	0.174 ± 0.075	0.138 ± 0.030
MalE-5	Alexa546-Alexa647	apo	0.326 ± 0.012	0.266 ± 0.013	0.295 ± 0.009
MalE-5	Atto532-Atto643	apo	0.232 ± 0.009	0.191 ± 0.009	0.211 ± 0.006
MalE-5	Alexa488-Alexa647	apo	0.135 ± 0.012	0.190 ± 0.042	0.160 ± 0.019
MalE-5	Alexa546-Alexa647	holo	0.330 ± 0.011	0.268 ± 0.040	0.297 ± 0.023
MalE-5	Atto532-Atto643	holo	0.232 ± 0.001	0.188 ± 0.018	0.209 ± 0.010
MalE-5	Alexa488-Alexa647	holo	0.120 ± 0.003	0.171 ± 0.089	0.143 ± 0.037
U2AF	Alexa546-Alexa647	apo	0.305 ± 0.041	0.261 ± 0.044	0.282 ± 0.030
U2AF	Atto532-Atto643	apo	0.229 ± 0.034	0.121 ± 0.010	0.166 ± 0.014
U2AF	Alexa488-Alexa647	apo	0.168 ± 0.045	0.253 ± 0.064	0.207 ± 0.038
U2AF	Alexa546-Alexa647	holo	0.332 ± 0.054	0.238 ± 0.012	0.281 ± 0.024
U2AF	Atto532-Atto643	holo	0.241 ± 0.013	0.106 ± 0.016	0.160 ± 0.012

U2AF	Alexa488-Alexa647	holo	0.150 ± 0.042	0.248 ± 0.062	0.193 ± 0.036

Supplementary Table 13. Computed distance uncertainties for MalE and U2AF65 samples with different dye combination using a "Diffusion with traps" (DWT) model and a "Wobbling in cone" (WIC) model. Uncertainties in the calculated distance due to uncertainities in the orientation factor, κ^2, are computed using the residual anisotropies of donor and acceptor fluorophores for a given FRET pair. For filtering out dye combinatins with specific sticking interactions, a threshold of 10% from the DWT model in the distance uncertainty was used. Details on uncertainty calculations using the DWT model[27] can be found in Supplementary Note 10. For details of the WIC model, see ref **23.**

Sample	Dye combination	state	$\Delta R_{app}(\kappa^2)$ [%] DWT model	$\Delta R_{app}(\kappa^2)$ [%] WIC model
MalE-1	Alexa546-Alexa647	apo	14.5	16.1
MalE-1	Alexa546-Abb. STAR635P	apo	11.0	13.3
MalE-1	Atto532-Atto643	apo	5.4	8.7
MalE-1	Alexa546-Alexa647	holo	14.8	14.9
MalE-1	Alexa546-Abb. STAR635P	holo	11.8	13.1
MalE-1	Atto532-Atto643	holo	6.2	9.0
MalE-2	Alexa546-Alexa647	apo	9.1	9.6
MalE-2	Atto532-Atto643	apo	5.2	7.2
MalE-2	Alexa488-Alexa647	apo	4.6	7.4
MalE-2	Alexa546-Alexa647	holo	8.2	9.6
MalE-2	Atto532-Atto643	holo	4.8	7.6
MalE-2	Alexa488-Alexa647	holo	4.5	7.5
MalE-3	Alexa546-Alexa647	apo	11.4	9.6
MalE-3	Atto532-Atto643	apo	3.3	5.2
MalE-3	Alexa488-Alexa647	apo	5.5	7.3
MalE-3	Alexa546-Alexa647	holo	9.4	8.5
MalE-3	Atto532-Atto643	holo	3.6	5.3
MalE-3	Alexa488-Alexa647	holo	5.8	7.4
MalE-4	Alexa546-Alexa647	apo	9.9	12.5
MalE-4	Atto532-Atto643	apo	3.7	7.0
MalE-4	Alexa488-Alexa647	apo	3.9	7.2
MalE-4	Alexa546-Alexa647	holo	13.3	14.1
MalE-4	Atto532-Atto643	holo	4.2	7.4
MalE-4	Alexa488-Alexa647	holo	4.0	7.0
MalE-5	Alexa546-Alexa647	apo	12.7	14.3
MalE-5	Atto532-Atto643	apo	7.1	9.7
MalE-5	Alexa488-Alexa647	apo	4.5	7.8
MalE-5	Alexa546-Alexa647	holo	16.2	14.4
MalE-5	Atto532-Atto643	holo	8.5	9.6
MalE-5	Alexa488-Alexa647	holo	4.3	7.2
U2AF	Alexa546-Alexa647	apo	14.3	13.4
U2AF	Atto532-Atto643	apo	5.9	8.2
U2AF	Alexa488-Alexa647	apo	8.1	9.6

U2AF	Alexa546-Alexa647	holo	11.7	13.5
U2AF	Atto532-Atto643	holo	5.3	8.1
U2AF	Alexa488-Alexa647	holo	5.9	9.1

Supplementary Table 14. Analysis of the U2AF65 dynamics. Results collected from 5 participating labs on the dynamics of U2AF65 labeled with the Atto532-Atto643 dye-pair. Timescales and rates were derived by dynamic PDA, filtered-FCS, and FRET-FCS as reported by the various groups. R_1 and R_2 are the distances corresponding to the compact and open states respectively for extracting the interconversion rates by dynamic PDA. Interconversion rates from closed to open and open to closed are denoted as k_{12} and k_{21} with their respective relaxation time τ_R. See Supplementary Note 18 for more details on the dynamics of U2AF65. The two extracted relaxation times from FRET-FCS and fFCS are denoted as t_{R1} and t_{R2}. See Supplementary Note 17 for more details on the filtered-FCS analysis. Average values of the rates and relaxation times are given in the last row.

Lab#	**Method**			**Apo**					**Holo**				
		R_1 (Å)	R_2 (Å)	t_{R1} (µs)	t_{R2} (µs)	k_{12} (ms^{-1})	k_{21} (ms^{-1})	τ_R (ms)	t_{R1} (µs)	t_{R2} (µs)	k_{12} (ms^{-1})	k_{21} (ms^{-1})	τ_R (ms)
1	Filtered-FCS			200	10								
	Dynamic PDA	38	59			0.43	0.07	2			0.52	0.14	1.42
2	Filtered-FCS			320	6				320	6			
	FRET-FCS*			32	2				321				
	Dynamic PDA	37	60			2.3	0.39	0.37			2.05	1.29	0.29
8	Dynamic PDA	38	56								0.62	0.18	1.25
11	Dynamic PDA	38	56								0.63	0.83	0.68
14	Filtered-FCS			370	12								
	Dynamic PDA	37	60			2.3	4.6	0.14	617	15	0.47	0.15	1.6
Average				296	9.3	1.67	1.68	0.83			0.85	0.51	1.05

* Only Lab#2 performed the FRET-FCS analysis and is not considered for average values for filtered-FCS.

Supplementary Table 15: The global dynamic photon distribution analysis (PDA) of apo and holo U2AF65 labeled with the Atto532-Atto643 dye pair. Parameters of the global dynamic PDA model are given, which is composed of a two-state dynamic system (apo ensemble) and two static states (holo and low-FRET) as described in the main text and Supplementary Note 18. The fit was performed globally over the apo and holo measurements using time windows of 0.5, 1.0, 1.5 and 2.0 ms, respectively. See Supplementary Figure 14 for an overview of the fits. The global reduced chi-square of the fit was 1.69. The analysis was performed using correction factors of: $\gamma = 0.59$ for the detection correction factor, a direct excitation correction of $\delta = 0.024$, a crosstalk value of $\alpha = 0.02$ and a Förster radius of 59 Å. Background count rates were 0.82 kHz and 0.28 kHz in the donor and FRET detection channels respectively. Fitting was performed using the *PDAFit* module of the PAM software package[8] .Errors were approximated from the covariance matrix given from the fit routine.

			Sample	
Population		**Parameter**	**apo**	**holo**
apo ensemble	compacted state	$R_{compact}$ [Å] $\sigma_{compact}$ [Å]	37.4 ± 0.1 2.1 ± 0.4	
	detached ensemble	$R_{detacted}$ [Å] $\sigma_{detached}$ [Å]	50.0 ± 0.1 4.5 ± 0.2	
	kinetic rates	$k_{c->d}$ [ms^{-1}] $k_{d->c}$ [ms^{-1}]	0.14 ± 0.11 0.013 ± 0.011	
	amplitude	A_{apo}	0.95 ± 0.02	0.35 ± 0.05
holo	open conformation	R_{open}[Å] σ_{open} [Å]	60.8 ± 0.1 2.9 ± 0.2	
	amplitude	A_{holo}	0.03 ± 0.01	0.60 ± 0.03
low-FRET	apparent distance	R_{LF}[Å] σ_{LF} [Å]	73.9 ± 4 7.2 ± 1.4	
	amplitude	A_{LF}	0.02 ± 0.01	0.05 ± 0.03

Supplementary Table 16. Correction factors obtained after reanalysis of the U2AF65 datasets from 8 different labs. Raw datasets were collected from 8 different labs for U2AF65 labeled with Atto532-Atto643. The reanalysis procedure is described in the Supplementary Note 4.

Lab#	***α***	***δ***	***γ***	***β***
1	0.02	0.06	0.73	0.72
2	0.04	0.20	1.26	0.22
3	0.03	0.06	0.98	0.77
4 *	0.03	0.08	0.98	1.15
7 **	0.03	0.05	1	1.22
8	0.03	0.05	1.15	0.80
11	0.03	0.07	0.65	1.21
14	0.03	0.05	0.92	1.00

* In these measurements, we observed a large contribution of the subpopulation with quenched acceptor lifetime (holo-high FRET) for double-labeled molecules. Hence, one population from apo and only one population from holo (holo-low FRET) conformation was used to calculate the γ globally.

** For these set of measurements, the holo measurement could not be analyzed. Hence, the given γ value was used.

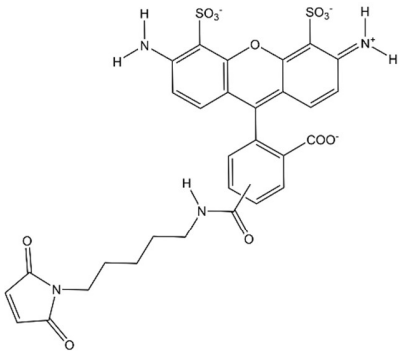
Supplementary Table 17. Dynamic photon distribution analysis (PDA) of apo and holo U2AF65 labeled with Alexa546-Alexa647 dye-pair. Parameters of the global PDA model are given, which is composed of a two-state dynamic system (apo ensemble) and two static states (holo and low-FRET) as described in the Supplementary Note 18. The fit was performed globally over the apo and holo measurements using time windows of 0.5, 1.0, and 1.5 ms respectively. See Supplementary Figure 17 for an overview of the fits. The analysis was performed using a detection correction factor of $\gamma = 0.32$, a direct excitation correction of $\delta = 0.10$, a crosstalk value of $\alpha = 0.036$ and a Förster radius of 65 Å. Background count rates were 0.918 kHz and 0.281 kHz in the donor and FRET detection channels respectively. Fitting was performed using the *PDAFit* module of the PAM software package[8].

			Sample	
Population		**Parameter**	**apo**	**holo**
apo ensemble	compacted state	$R_{compact}$ [Å]	35.4	
		$\sigma_{compact}$ [Å]	3.5	
	detached ensemble	$R_{detacted}$ [Å]	50.3	
		$\sigma_{detached}$ [Å]	5.0	
	kinetic rates	$k_{c->d}$ [ms^{-1}]	0.16	
		$k_{d->c}$ [ms^{-1}]	0.014	
	amplitude	A_{apo}	0.97	0.36
holo	open conformation	R_{open}[Å]	-	62.0
		σ_{open} [Å]	-	2.9
	amplitude	A_{holo}	-	0.59
low-FRET	apparent distance	R_{LF}[Å]	72.5	
		σ_{LF} [Å]	5.0	
	amplitude	A_{LF}	0.03	0.05

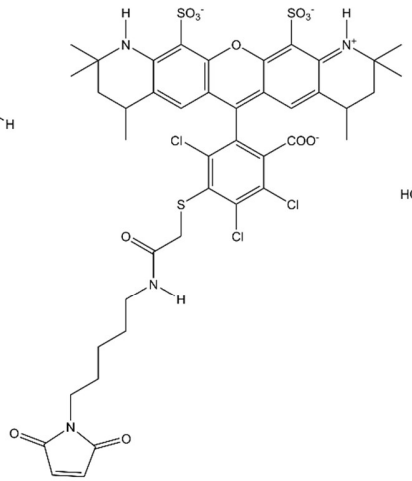
Supplementary Table 18. Detailed information on dye maleimides. Dye-maleimide conjugates were purchased from the listed companies. Structural formulas of dyes are illustrated below, which exception of Atto643 where chemical structure of the dye is unknown. For dyes Atto532 and Abberior STAR635P, linkage length to maleimide functional group is not indicated by producer. Those dyes are therefore illustrated without maleimide modification.

Fluorophore	Company	Catalog Number
Alexa Fluor 488 C5 Maleimide	ThermoFisher Scientific	A10254
Alexa Fluor 546 C5 Maleimide	ThermoFisher Scientific	A10258
Alexa Fluor 647 C2 Maleimide	ThermoFisher Scientific	A20347
Atto532 Maleimide	ATTO-TEC	AD 532-45
Atto643 Maleimide	ATTO-TEC	AD 643
Abberior STAR635P Maleimide	Abberior	ST635P

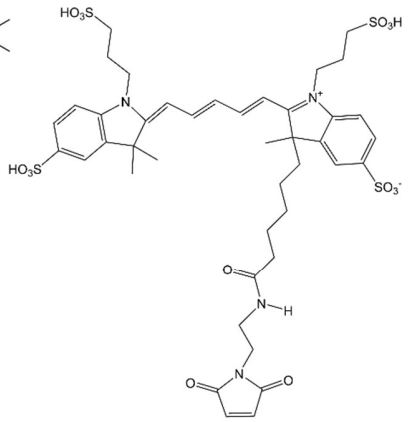
Alexa Fluor 488 C5 Maleimide



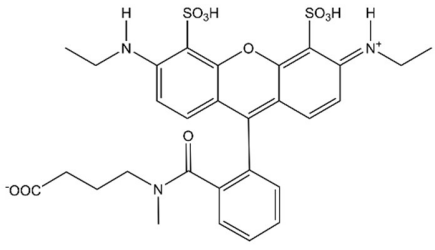
Alexa Fluor 546 C5 Maleimide



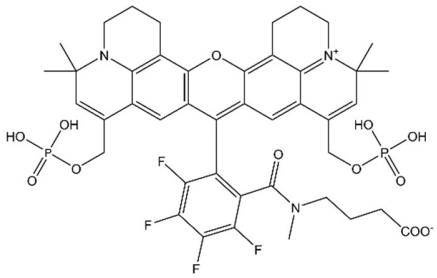
Alexa Fluor 647 C2 Maleimide



Atto 532



Abberior STAR635P



Supplementary Table 19. Parameters used for the AV and ACV calculations. The fraction of trapped dye is computed from the fundamental and residual anisotropy as $x_{\text{trapped}} = r_{\infty,tr}/r_0$ (see Supplementary Table 9 and Supplementary Note 9, Eqn. 9.8).

	Dye species	
	Alexa Fluor 546	**Alexa Fluor 647**
Dye parameters		
L_{length} / Å	20.5	21.0
L_{width} / Å	4.5	4.5
$R_{\text{dye,1}}$ / Å	5.0	11.0
$R_{\text{dye,2}}$ / Å	4.5	4.7
$R_{\text{dye,3}}$ / Å	1.5	1.5
AV parameters		
Grid step / Å	0.9	
Allowed sphere radius / Å	1.0	
ACV parameters		
Contact volume thickness / Å	3.0	
Mutant	**Fraction of trapped dye (apo/holo)**	
29	0.87/0.83	0.33/0.36
352	0.72/0.73	0.68/0.74
87	0.54/0.62	0.44/0.44
186	0.38/0.42	0.22/0.22
134	0.90/0.90	0.45/0.43
34-205*	0.91/0.92	0.54/0.57
36-205*	0.88/0.89	0.72/0.78

*Anisotropy measurements of single mutants were not performed. Hence, trapped fractions were obtained from smFRET measurements and, as such, are an average of the two labeling positions due to the stochastic labeling of the protein.

Supplementary References:

1. Peulen, T. O., Opanasyuk, O. & Seidel, C. A. M. Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. J. Phys. Chem. B 121, 8211–8241 (2017).

2. Haas, E., Ephraim-Katchalski-Katzir & Steinberg, I. Z. Effect of the Orientation of Donor and Acceptor on the Probability of Energy Transfer Involving Electronic Transitions of Mixed Polarization. Biochemistry 17, 5064–5070 (1978).

3. Kong, X., Nir, E., Hamadani, K. & Weiss, S. Photobleaching pathways in single-molecule FRET experiments. J. Am. Chem. Soc. 129, 4643–4654 (2007).

4. Kudryavtsev, V. et al. Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. ChemPhysChem 13, 1060–1078 (2012).

5. Eggeling, C. et al. Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. Journal of Biotechnology 86, 163-180 (2001).

6. Lerner, E. et al. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. eLife 10, e60416 (2021).

7. De Boer, M. et al. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. eLife 8, e44652 (2019).

8. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. Biophys. J. 114, 1518–1528 (2018).

9. Tomov, T. E. et al. Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. Biophys. J. 102, 1163–1173 (2012).

10. Lee, N. K. et al. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. Biophys. J. 88, 2939–2953 (2005).

11. Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. Nat. Methods 15, 669–676 (2018).

12. Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying molecular dynamics in single-molecule fret experiments with burst variance analysis. Biophys. J. 100, 1568–1577 (2011).

13. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. M. Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. J. Phys. Chem. B 114, 7983–7995 (2010).

14. Barth, A. et al. Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. J. Chem. Phys. 156, 141501 (2022).

15. Foerster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. Ann. Phys. 437, 55–75 (1948).

16. ur Rehman Malik, N. & Sadarangani, C. Brushless doubly-fed induction machine with rotating power electronic converter for wind power applications. 2011 Int. Conf. Electr. Mach. Syst. 1–6 (2011).

17. Wiederschain, G. Y. The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies. Biochem. 76, 1276–1276 (2011).

18. Voith von Voithenberg, L. & Lamb, D. C. Single Pair Förster Resonance Energy Transfer: A Versatile Tool To Investigate Protein Conformational Dynamics. BioEssays 40, 1–14 (2018).

19. Spectra, E. Alexa Fluor ® Dyes — Across the Spectrum. 1–3 (2014).

20. Schaflfer, J. et al. Identification of Single Molecules in Aqueous Solution by Time-Resolved Fluorescence Anisotropy. J. Phys. Chem. A 103, 331–336 (1999).

21. Perrin, F. Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l'etat excité. J. Phys. le Radium 7, 390–401 (1926).

22. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection. Methods Enzymol. 475, 455–514 (2010).

23. Sindbert, S. et al. Accurate distance determination of nucleic acids via Förster resonance energy transfer: Implications of dye Linker length and rigidity. J. Am. Chem. Soc. 133, 2463–2480 (2011).

24. Koshioka, M., Sasaki, K. & Masuhara, H. Time-Dependent Fluorescence Depolarization Analysis in Three-Dimensional Microspectroscopy. Appl. Spectrosc. 49, 224–228 (1995).

25. Dale, R. E. & Eisinger, J. Intramolecular distances determined by energy transfer. Dependence on orientational freedom of donor and acceptor. Biopolymers 13, 1573–1605 (1974).

26. Ivanov, V., Li, M. & Mizuuchi, K. Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. Biophys. J. 97, 922–929 (2009).

27. Kalinin, S., Fulle, S., Hanke, C.A., Peulen, T. O., Sindbert, S., Felekyan, S., Kühnemuth, R., Gohlke, H., Seidel, C. A. M. Diffusion with traps: experiment, simulation, and theory to describe the dynamics of flexibly linked fluorophores in biomolecular FRET. Prep.

28. D.A. Case et al. AMBER 2018, University of California. University of California, San Francisco (2018).

29. Maier, J. A. et al. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theory Comput. 11, 3696–3713 (2015).

30. Horn, H. W. et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J. Chem. Phys. 120, 9665–9678 (2004).

31. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. J. Chem. Theory Comput. 11, 1864–1874 (2015).

32. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23, 327–341 (1977).

33. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092 (1993).

34. McGibbon, R. T. et al. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. Biophys. J. 109, 1528–1532 (2015).

35. Gopich, I. V. & Szabo, A. Single-molecule FRET with diffusion and conformational dynamics. J. Phys. Chem. B 111, 12925–12932 (2007).

36. Zander, C. et al. Detection and characterization of single molecules in aqueous solution. Appl. Phys. B Lasers Opt. 63, 517–523 (1996).

37. Hall, P. & Selinger, B. Better estimates of exponential decay parameters. Journal of Physical Chemistry 85, 2941-2946 (1981).

38. Livesey, A. K. & Skilling, J. Maximum entropy theory. Acta Crystallogr. Sect. A 41, 113–122 (1985).

39. Brochon, J. C. Maximum entropy method of data analysis in time-resolved spectroscopy. Methods Enzymol. 240, 262–311 (1994).

40. Skilling, J. & Bryan, R. K. Maximum entropy image reconstruction: general algorithm. Mon. Not. R. Astron. Soc. 211, 111–124 (1984).

41. Vinogradov, S. A. & Wilson, D. F. Recursive maximum entropy algorithm and its application to the luminescence lifetime distribution recovery. Appl. Spectrosc. 54, 849–855 (2000).

42. Huang, J. R. et al. Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: A combined NMR and SAXS study. J. Am. Chem. Soc. 136, 7068–7076 (2014).

43. Kalinin, S. et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat. Methods 9, 1218–1225 (2012).

44. Bowman, A. W. & Azzalini, A. Computational aspects of nonparametric smoothing with illustrations from the sm library. Comput. Stat. Data Anal. 42, 545–560 (2003).

45. Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A. & Seidel, C. A. M. Filtered FCS: Species auto- and cross-correlation functions highlight binding and dynamics in biomolecules. ChemPhysChem 13, 1036–1053 (2012).

46. Barth, A. et al. Dynamic interactions of type I cohesin modules fine-tune the structure of the cellulosome of Clostridium thermocellum. Proc. Natl. Acad. Sci. U. S. A. 115, E11274–E11283 (2018).

47. Von Voithenberg, L. V. et al. Recognition of the 3′ splice site RNA by the U2AF heterodimer involves a dynamic population shift. Proc. Natl. Acad. Sci. U. S. A. 113, E7169–E7175 (2016).

48. Margittai, M. et al. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. Proc. Natl. Acad. Sci. U. S. A. 100, 15516–15521 (2003).

49. Torres, T. & Levitus, M. Measuring conformational dynamics: A new FCS-FRET approach. J. Phys. Chem. B 111, 7392–7400 (2007).

50. Felekyan, S., Sanabria, H., Kalinin, S., Kühnemuth, R. & Seidel, C. A. M. Analyzing Förster resonance energy transfer with fluctuation algorithms. Methods Enzymol. 519, 39–85 (2013).

51. Antonik, M., Felekyan, S., Gaiduk, A. & Seidel, C. A. M. Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. J. Phys. Chem. B 110, 6970–6978 (2006).

52. Kalinin, S., Felekyan, S., Antonik, M. & Seidel, C. A. M. Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer. J. Phys. Chem. B 111, 10253–10262 (2007).

53. Sánchez-Rico, C., Voith von Voithenberg, L., Warner, L., Lamb, D. C. & Sattler, M. Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by spFRET and NMR Spectroscopy. Chem. - A Eur. J. 23, 14267–14277 (2017).

54. Barth, A., Voith Von Voithenberg, L. & Lamb, D. C. Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. J. Phys. Chem. B 123, 6901–6916 (2019).

55. Fries, J. R., Brand, L., Eggeling, C., Köllner, M. & Seidel, C. A. M. Quantitative identification of different single molecules by selective time-resolved confocal fluorescence spectroscopy. J. Phys. Chem. A 102, 6601–6613 (1998).

56. Gouridis, G. et al. Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. Nat. Struct. Mol. Biol. 22, 57–64 (2015).

57. de Boer, M., Gouridis, G., Muthahari, Y. A. & Cordes, T. Single-Molecule Observation of Ligand Binding and Conformational Changes in FeuA. Biophys. J. 117, 1642–1654 (2019).

58. Wolf, S. et al. Hierarchical dynamics in allostery following ATP hydrolysis monitored by single molecule FRET measurements and MD simulations. Chem. Sci. 12, 3350–3359 (2021).

59. Ambrose, B. et al. The smfBox is an open-source platform for single-molecule FRET. Nat. Commun. 11, 5641 (2020).

60. Schwarz, M. et al. Single-molecule nucleosome remodeling by INO80 and effects of histone tails. FEBS Lett. 592, 318–331 (2018).

61. Tan, P. S. & Lemke, E. A. Probing Differential Binding Mechanisms of Phenylalanine-Glycine-Rich Nucleoporins by Single-Molecule FRET. in Methods in Enzymology 611, 327–346 (2018).

62. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. Biophys. J. 89, 3508–3522 (2005).

63. Robb, N. C. et al. The transcription bubble of the RNA polymerase-promoter open complex exhibits conformational heterogeneity and millisecond-scale dynamics: Implications for transcription start-site selection. J. Mol. Biol. 425, 875–885 (2013).

64. Santoso, Y. et al. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. Proc. Natl. Acad. Sci. U. S. A. 107, 715–720 (2010).

65. Santoso, Y., Torella, J. P. & Kapanidis, A. N. Characterizing single-molecule FRET dynamics with probability distribution analysis. ChemPhysChem 11, 2209–2219 (2010).

66. Robb, N. C. et al. Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA. Nucleic Acids Res. 44, 10304–10315 (2016).

67. Tomescu, A. I., Robb, N. C., Hengrung, N., Fodor, E. & Kapanidis, A. N. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. Proc. Natl. Acad. Sci. 111, E3335-3342 (2014).

68. Kapanidis, A. N. et al. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. Proc. Natl. Acad. Sci. U. S. A. 101, 8936–8941 (2004).

69. Kapanidis, A. N. et al. Alternating-laser excitation of single molecules. Acc. Chem. Res. 38, 523–533 (2005).

70. Krainer, G. et al. Ultrafast Protein Folding in Membrane-Mimetic Environments. J. Mol. Biol. 430, 554–564 (2018).

71. Hartmann, A., Krainer, G., Keller, S. & Schlierf, M. Quantification of Millisecond Protein-Folding Dynamics in Membrane-Mimetic Environments by Single-Molecule Förster Resonance Energy Transfer Spectroscopy. Anal. Chem. 87, 11224–11232 (2015).

72. Hartmann, A. Observing Biomolecular Dynamics from Nanoseconds to Hours with Single-Molecule Fluorescence Spectroscopy. 139 (2017).

73. Gouge, J. et al. Molecular mechanisms of Bdp1 in TFIIIB assembly and RNA polymerase III transcription initiation. Nat. Commun. 8, 130 (2017).

74. Kahra, D. et al. Conformational plasticity and dynamics in the generic protein folding catalyst SlyD unraveled by single-molecule FRET. J. Mol. Biol. 411, 781–790 (2011).

75. Olofsson, L. & Margeat, E. Pulsed interleaved excitation fluorescence spectroscopy with a supercontinuum source. Opt. Express 21, 3370 (2013).

76. Fuertes, G. et al. Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. Proc. Natl. Acad. Sci. U. S. A. 114, E6342–E6351 (2017).

77. Nikaido, H. Maltose transport system of Escherichia coli: An ABC-type transporter. FEBS Lett. 346, 55–58 (1994).

78. McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophys. J. 91, 1941–1951 (2006).

79. Jerabek-Willemsen, M. et al. MicroScale Thermophoresis: Interaction analysis and beyond. J. Mol. Struct. 1077, 101–113 (2014).

80. Widengren, J. & Schwille, P. Characterization of photoinduced isomerization and back-isomerization of the cyanine dye cy5 by fluorescence correlation spectroscopy. J. Phys. Chem. A 104, 6416–6428 (2000).

81. Petrášek, Z. & Schwille, P. Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. Biophys. J. 94, 1437–1448 (2008).

7.2 An optofluidic antenna for enhancing fluorescence detection

An optofluidic antenna for enhancing fluorescence detection

Luis Morales-Inostroza,[1,2,3,*] Julian Folz,[4,*] Ralf Kühnemuth,[4] Suren Felekyan,[4] Franz-Ferdinand Wieser,[1,2,3] Claus A.M. Seidel,[4,†] Stephan Götzinger,[3,1,5] and Vahid Sandoghdar[1,3,‡]

[1] *Max Planck Institute for the Science of Light, 91058 Erlangen, Germany*
[2] *Max-Planck-Zentrum für Physik und Medizin, 91058 Erlangen, Germany*
[3] *Department of Physics, Friedrich-Alexander-Universität of Erlangen-Nürnberg, 91058 Erlangen, Germany*
[4] *Chair for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*
[5] *Graduate School in Advanced Optical Technologies (SAOT), FAU, D-91052 Erlangen, Germany*

(Dated: July 7, 2022)

Many single-molecule optical studies are performed in fluidic environments, e.g., to avoid unwanted consequences of protein immobilization. The inherent diffusion of molecules in this arrangement limits the observation time and the number of collected photons. It, thus, follows that studies of processes with both fast and slow dynamics are compromised. Here, we introduce an optofluidic dielectric antenna (OFA), which enhances the fluorescence brightness by up to a factor of 5 per passage, leads to about 7 times more frequent returns to the observation volume, and results in about 7.5 times longer diffusion times within one passage. We use single-molecule multi-parameter fluorescence detection (sm-MFD), fluorescence correlation spectroscopy (FCS) and Fluorescence Resonance Energy Transfer (FRET) measurements to characterize the basic features of our OFAs. We then showcase the advantages of our antenna by examining both slow (1 ms) and fast (100 μs) dynamical behavior of Holliday Junctions. Their ease of implementation and compatibility with various microscopy modalities makes OFAs broadly applicable to a large variety of studies.

Keywords: Collection efficiency, Dielectric antenna, FCS

Since its debut in the early 1990s, single-emitter fluorescence detection has found applications in many fields, ranging from biophysics to quantum optics [1–5]. The quality and success rate of single-emitter studies, however, can vary and are often limited by the shot noise in the number of detected photons. It follows that sensitive fluorescence measurements are particularly challenging under fluidic conditions, where diffusion restricts the observation time. In addition to limiting the signal-to-noise ratio (SNR), a short measurement time also hampers the investigation of slow phenomena. Single-molecule studies in fluids also confront a challenge for measurements of fast (sub-ms) phenomena because the finite lifetime of the excited state imposes an intrinsic upper bound on the fluorophore emission rate, thus, limiting the minimal integration time. To extend the observation time without jeopardizing contact with surfaces, various trapping schemes have been explored [6–8]. It is also possible to enhance the emission rate by employing optical micro-resonators and plasmonic nano-antennas although these usually require precise positioning and orientation of the fluorophore in a confined mode and are only effective in a narrow range of wavelengths [9].

In the past decade, we have proposed and demonstrated a versatile planar antenna that provides collection efficiencies as high as 99% over a very broad spectral domain [10–13]. In this work, we show how this concept can be implemented in fluidic arrangements for enhancing the collected number of photons per unit time. Furthermore, we show that the geometry of such an optofluidic dielectric antenna (OFA) extends the observation time by up to 7.5 times. We demonstrate the virtues of this new technology with two concrete examples. The experimental ease of operation of OFAs provides a powerful platform for performing novel fluorescence measurements.

I. OPTOFLUIDIC ANTENNA: THEORY AND DESIGN

The underlying physics of a dielectric planar antenna is based on reshaping the radiation pattern of an emitter by means of a stratified dielectric structure that acts as a leaky quasi-waveguide [10]. The operation principle follows the rule of thumb that $n_1 > n_2 > n_3$, where the emitter is placed within a medium with refractive index n_2, sandwiched between a medium of n_3 and a substrate with n_1. It then follows that the emitter radiation is efficiently channeled into the higher index medium and onto a detector, e.g., through a microscope objective. Medium 2 is chosen to be a few wavelengths thick although the performance of the device is quite tolerant of small deviations in the medium thickness and of the exact position of the emitter in it [11]. Another noteworthy feature of the planar antennas is a large spec-

* These two authors contributed equally
† Correspondence email address: cseidel@hhu.de
‡ Correspondence email address: vahid.sandoghdar@mpl.mpg.de

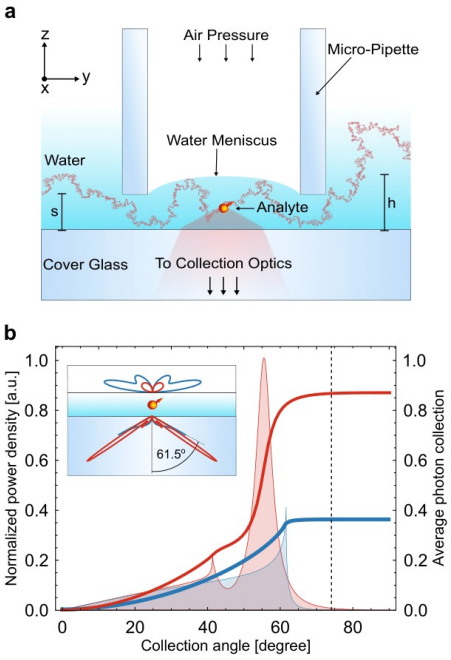


Figure 1. **Optofluidic antenna designed for** 86% **photon collection efficiency from a randomly oriented dipole**. **a,** Sketch of a single molecule diffusing through an OFA (red dashed curve). The refractive indices of the different materials are $n_1 = 1.0$ (air), $n_2 = 1.33$ (water) and $n_3 = 1.517$ (glass). The water layer thickness is set to $h = 500\,\mathrm{nm}$, the distance between the micropipette and the cover glass is $s = 100\,\mathrm{nm}$, the micro-pipette inner diameter and wall thickness are $14\,\mu\mathrm{m}$ and $3\,\mu\mathrm{m}$, respectively. **b,** Calculated emitted power density and collection efficiency averaged for random dipole orientations as a function of the collection angle inside the OFA (red) and in an extended solution (blue). The vertical black dashed line indicates the maximum collection angle of the microscope objective. Inset: radiated power density in polar coordinates.

tral bandwidth spanning over more than three hundred nanometers in wavelength [11]. Experimental demonstrations of planar antennas have previously focused on solid-state emitters and applications in quantum optics [10, 12, 13]. In this work, we introduce the opto*fluidic* antenna (OFA).

A technical challenge in the realization of an OFA is to maintain the thickness of the fluidic medium 2 in the range of one micrometer or less. While nanofluidic chips can easily fulfil this condition between two glass or polymer media [6], bordering a medium of lower refractive index is nontrivial for aqueous media of this thickness. To produce this unconventional arrangement, we have devised a setup that allows one to confine a fluid by introducing a body of gas in a controlled and local fashion.

The schematics of our OFA is shown in Fig. 1**a**. A heat-pulled micropipette with an inner diameter in the range of $10\,\mu\mathrm{m}$ is dipped in a liquid and placed close to the interface with the underlying substrate. By chemically modifying the inner wall surface of the micropipette (see Methods), we prevent water from entering, thus accommodating a gaseous upper medium 3 over the micropipette opening. The height of the resulting fluidic channel is precisely tunable using a piezo actuator. Additionally, a syringe connected to the non-tapered end of the pipette (not shown in Fig. 1**a**) is used to control the shape of the water meniscus inside the micropipette by adjusting the gas pressure. We use a standard cover glass as a substrate in combination with an oil-immersion microscope objective. Alternatively, a thin layer of a dielectric (in our case of MgF_2) can be applied to the substrate as a spacer to prevent the analyte from undergoing evanescent coupling to the underlying substrate [10]. A decisive advantage of our OFA design is that it can easily be implemented in existing inverted microscopes.

Figure 1**b** displays the number of collected photons emitted by a dipole inside the OFA (red) and a dipole in an extended solution (blue) as a function of the opening angle. As the right-hand vertical axis shows, we can achieve an average photon collection efficiency that is greater than 86% for a randomly oriented dipole in a fluid. We note that enhancements resulting from evanescent coupling are not desirable because they imply vicinity to surfaces and are strongly distance dependence. The enhanced collection efficiency provided by the OFA amounts to nearly 2.2 fold compared to the case of an extended solution. In addition to the effects of a planar dielectric antenna [10, 12, 13], the quasi-two-dimensional geometry of OFA also brings about other significant advantages, which we discuss in the next section.

II. FLUORESCENCE CORRELATION SPECTROSCOPY

OFA can be used in a number of applications where higher photon budgets are advantageous. To demonstrate the main features of an OFA, we now discuss our measurements on fluorescence correlation spectroscopy (FCS) of diffusing dye molecules in solution. In FCS, one usually analyzes the fluorescence signal that is generated as particles diffuse through the focus of a laser beam (see SI section VI D). Application of a pinhole in the detection path restricts the origin of the signal to an

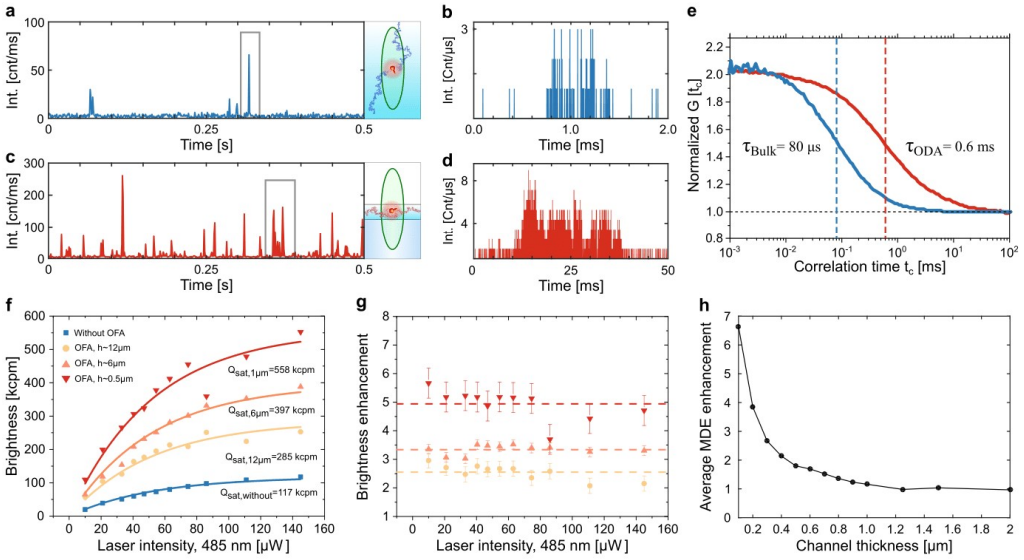


Figure 2. **Fluorescence correlation spectroscopy in OFA. a,** FCS time trace recorded in a standard arrangement in an open solution. **b,** Zoom in an exemplary burst from (a). **c,** A time trace recorded inside the OFA for a water layer thickness of 0.5 μm. **d,** Zoom in an exemplary bursts obtained from analytes diffusing inside the OFA. Sketch in (a) and (c) illustrate the diffusion in bulk and inside the OFA, respectively. **e,** Typical cross-correlation curves obtained when the analytes diffuse in an open solution (blue curve) and inside the OFA (red curve). **f,** Power-dependent brightness in confocal measurements of freely diffusing Rhodamin110 molecules obtained via FCS at different heights h for the antenna. Each data point represents a 10 min measurement where the auto-correlation (AC) function was fitted using an equation with a bunching term and a 3-dimensional Gaussian shaped confocal volume. The brightness is stated as the ratio of the obtained signal divided by the number of molecules in a bright state (triplet corrected) within the focus. The model fitted a global value of 64 μW. **g** Brightness enhancement as a function of the excitation power. The color code follows the labels in (f). The dashed horizontal lines indicate the average brightness enhancement obtained for each thickness of the water layer. **h,** Relative enhancement in the molecular detection efficiency as a function of the channel thickness. The lateral extension (w_0 =1 μm) is extracted from the calibration of the experimental setup, whereby w_0 = is defined as the width, where the intensity decays to $1/e^2$ of its maximum. The pinhole size was set to coincide with the lateral extension of the beam width in the focal plane. Inset: Illustration of diffusion inside the OFA and in bulk solution. The point-spread function of the emitter was simulated according to the work of Mortensen et al. [14].

elliptical volume. The detection efficiency is commonly characterized by the molecular detection efficiency function (MDE) [15, 16]. The area under a fluorescence burst indicating the cumulative signal emitted during one passage can, thus, be defined as "brightness".

In Fig. 2**a**, we display an example of fluorescence bursts recorded over a period of 500 ms from Rhodamin 110 (Rh110) dye molecules diffusing in water in the absence of an OFA. Figure 2**b** displays a close-up of a strong burst spread over about 1 ms, as expected from the passage of a single molecule. Since each molecule diffuses through a different stochastic path, bursts can vary in duration and in the number of photons. Nevertheless, a visual comparison of the data in Fig. 2**a,b** with those presented in Fig. 2**c,d** for molecules in an OFA indicates that 1) the number of photons per unit of time is larger when using the OFA, 2) the bursts last longer, and 3) the overall number of bursts recorded with the OFA is considerably higher than in the case of diffusion in an extended solution. While the first effect is expected from a dielectric planar antenna, the two other observations have their origin in the microfluidic arrangement of the OFA.

To scrutinize the duration of bursts, in Fig. 2**e** we displays characteristic second-order autocorrelation functions of the fluorescence signal originating from the confocal spot of a microscope objective with a numerical aperture NA = 1.4. The observed correlation time of 80 μs in the blue curve suggests an average diffusion time of about 0.1 ms for a molecule in open geometry

if one models the observation volume as a 3D Gaussian shape with a lateral of about 0.5 μm [17]. The red curve in Fig. **2e** shows that the correlation time and, therefore, the passage time are substantially increased (here by 7.5 times) in the OFA arrangement. We attribute the slow diffusion to the affinity of dye molecules for the water-air interface. Indeed, studies on water micro-droplets in oil have shown that dye molecules preferably diffuse closer to the water-oil interface [18, 19]. In one scenario, it has been stated that a static electric field confines the diffusion of molecules by aligning their dipole moments at the interface [18–20]. Other experimental data indicate that the accumulation of molecules in water micro-droplets is independent of the molecular charge [20]. In the case of the interface between water and air, however, the exact composition and structure of the interface, e.g., its acidic-basic nature, remains unknown [21–23]. To explain longer diffusion times in an OFA, we introduced an interface residence time in Monte Carlo simulations. We find that a value of 40 μs upon each interface visit reproduces our experimental observations. We also investigated this effect experimentally by examining dye molecules with different charge states (see SI section VI E for details). However, a quantitative understanding of these phenomena is beyond the scope of our current work, but the architecture of the OFA offers an ideal platform for studying the effect of the water-air interface on molecular diffusion.

The axial confinement of molecular motion in an OFA forces passage through the maximum of the detection volume as determined by the pinhole. To examine the effect of phenomenon, we varies the distances of the micropipette end from the water-glass interface to achieve different channel thicknesses. In addition, at each condition we varied the excitation power to ensure that it does not affect the antenna performance. Figure **2f** presents the outcome for the conventional FCS arrangement (blue) and several intermediate thicknesses of the OFA. The solid lines fitted to the experimental data show that the fluorescence signal follows the expected saturation curve at higher powers. The data show a 5-fold enhancement in brightness for the most confined OFA configuration. Figure **2g** verifies that the brightness increase is independent of the excitation power. Considering that the collection efficiency is improved by 2.2 times due to the effect of the dielectric planar antenna, we conclude the OFA brings about an extra enhancement factor of 2.5.

To gain more insight into this observation, we also performed analytical calculations, whereby we solved for the emission pattern of a dye located at an arbitrary position under the condition that it passes through the pinhole in the detection path (see Methods and SI). In a standard FCS configuration, analytes in the buffer undergo Brownian motion in a large region surrounding the observation volume, so that most molecules that traverse the field of view do not contribute significantly to the signal. In an OFA, however, the confinement imposed by the thin fluidic layer ensures that all molecules that cross the laser beam contribute to the signal on the detector. Figure **2h** shows the resulting enhancement factor as compared to the case of an open liquid. We find that the molecular detection efficiency is indeed enhanced by 2.5 times if the channel height is about 300 nm,. For water layer thickness beyond $\approx 1\,\mu$m, the enhancement factor drops to 1.

To understand the third effect, namely a higher number of bursts in an OFA, we investigated the influence of the fluidic confinement on the diffusion behavior by performing Monte-Carlo simulations. The outcome shows that for a channel height of 0.5 μm, the *probability to return* to the observation volume is approximately 9 times larger in the quasi-2D geometry of the OFA than for the case of an open buffer. This agrees well with our experimental observation of up to about 7 times more bursts if count rates that are 3σ larger than the background are considered. More frequent passages allow shorter measurement times at a given concentration although it could also lead to crowding if the concentration is not adjusted accordingly.

Finally, we note that the stratified architecture of an OFA can also affect the distribution of the excitation laser intensity and thus the observation volume. To investigate this, we performed finite difference time-domain (FDTD) simulations. We find that the lateral extension of the observation volume is only modified by a few percent in the OFA geometry (see SI section VI A for details), whereas the axial dimension of the observation volume can be reduced in the OFA geometry, depending on the NA of the excitation beam. Moreover, one might wonder if the presence of the water-glass interface in an OFA changes the excited-state lifetime of the fluorophore. We, thus, measured the fluorescence lifetime with and without the antenna. We find that a value of $\tau = 4.05$ ns in an OFA and a lifetime of $\tau = 3.98$ ns in an open buffer coincide within the measurement error of 0.1 ns, indicating that a geometry-induced lifetime change is negligible.

We have shown that an OFA not only improves the collection efficiency of fluorescence emission but it also extends the diffusion time and thus the overall brightness. Both of these phenomena are highly desirable for a wide range of sensitive studies on 1) weakly emitting fluorophores, 2) fast events that would benefit from higher photon rates and thus shorter integration times, and 3) slow events that require longer observation times. In the next section, we present a case study, where an OFA is used to investigate dynamic processes.

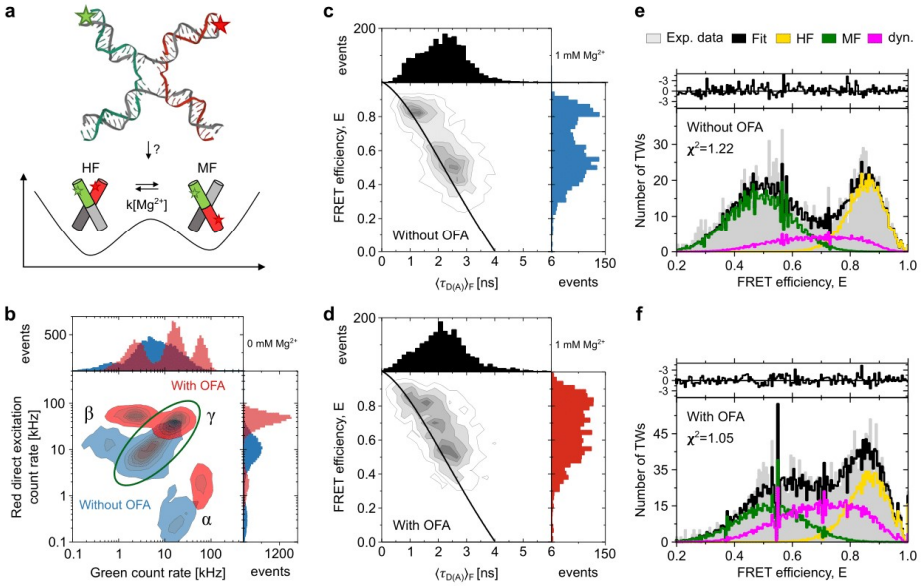


Figure 3. **Holliday junction single molecule multiparameter fluorescence detection (sm-MFD). a,** Sketch of the transition of the HJ between the two stacked conformations assigned by high FRET (HF) and middle FRET (MF) states. The labeled arms of the junction are shown in green for the donor and red for the acceptor. Between the two stable FRET states the molecule has to pass a planar cross state where the donor-acceptor distance is maximized, leading to a low FRET (LF) state. The rate $k[Mg^{2+}]$ between these two stable states depends on the concentration of Mg^{2+} in the buffer: it is high without Mg^{2+} and becomes lower when adding Mg^{2+}. **b,** 2D histograms showing the green and red count rates in direct excitation obtained from single molecule bursts with two 1D projections. Three populations can be identified and interpreted as donor only molecules (α), acceptor only molecules (β) and double-labeled molecules (γ). Only double labeled molecules were selected for the analysis using a stoichiometry and an ALEX-2CDE filter.**c,d** 2D histograms of sm-MFD measurements in a 1 mM Mg^{2+} buffer without (c) and with OFA (d). Only doubly-labeled molecules were selected. The black lines represent the static FRET lines. The measurement with the OFA shows more dynamical averaging in the region between HF and MF populations since the diffusion time is enlarged. **e,f** Corresponding PDA of the shown sm-measurements with/without the OFA. The shown histograms are based on a 2 ms TW. The analysis was done globally for TWs of 1, 2 and 3 ms. A two-state model of HJ was applied where the dynamic fraction between the HF and MF population was observed and extracted. The top panels in (e) and (f) show the residuals of the fit process to the experimental data. The HF static state is plotted in dark yellow, MF static state in dark green and the dynamic fraction of molecules in between these two states is shown in light pink. Resulting superposition of these 3 states is shown in black and compared to the experimentally obtained histogram shown in grey.

III. FLUORESCENCE RESONANCE ENERGY TRANSFER: HOLIDAY JUNCTION DYNAMICS

Holliday junctions (HJ) are four-way DNA junctions formed during DNA strand exchange of homologous duplexes and are fundamental in genetic recombination [24]. In buffer solution, HJs exist in two conformations with branches stacked in a parallel or anti-parallel fashion (see sketch in Fig.3**a**). The fluctuation between these two conformations has been the subject of many studies, and it was found that among other factors, it depends on the concentration of Mg^{2+}. We, therefore, used this system for a case study in which both slow and fast dynamic processes are manifested.

HJs were studies in a buffer solution and labeled with Alexa488 as donor (green channel) and Atto647N as acceptor (red channel) at strands *a* and *b*, respectively. The details of the sample preparation can be found in Methods and SI section VI G. We carefully characterized the internal HJ dynamics during diffusion inside the OFA to ensure that the antenna does not perturb them (see SI section VI J). We used single-molecule multi-parameter fluorescence detection

(sm-MFD) in a pulsed interleaved excitation (PIE) scheme [25] to achieve a polarization-dependent, time-correlated single-photon counting, providing high accuracy and precision. In this manner, we extract multiple parameters such as the fluorescence lifetime, FRET efficiency, stoichiometry, correlation amplitudes, and anisotropy values simultaneously [26]. Moreover, calibrated polarization measurements let us conclude that the OFA arrangement results in a g-factor of 1.2, defined as the ratio of parallel to perpendicular polarization for the detected light. Single-molecule measurements yield a residual anisotropy of r_∞ =0.06, indicating a free rotation and therefore unrestricted movement of the HJ and of the dye (see SI section VI G).

In Fig. 3**b**, we present the recorded count rates for the donor and acceptor channels in a 2D MFD-histogram as well as their corresponding 1D projections for experiments performed without (blue) and with an OFA (red). We identify three populations for the donor only (β), acceptor only (α) and doubly-labeled case (γ). It is evident that the histogram widths become narrower for both channels whereas the ratio of green and red count rates remains unchanged. This is because the shot noise could be reduced by collecting a larger number of photons. We verified that the antenna enhances the average photon detection rate by a factor of 4 for both donor and acceptor dyes. Even though the donor and acceptor emission wavelengths are spectrally separated by more than 150 nm, they experience the same enhancement in collection efficiency, a property that follows from the broadband nature of the antenna design in contrast to other structures such as microcavities or plasmonic nano-antennas. This feature advocates the use of OFAs for accurate FRET-derived distance measurements at Ångstrom precision.

To ensure that the OFA does not introduce systematic effects on the internal dynamics of the structure under study, we examined the influence of Mg^{2+} ions on the FRET efficiency E and the fluorescence lifetimes $\tau_{D(A)}$ of the donor in the presence of an acceptor molecule. In the absence of Mg^{2+}, HJ displays high relaxation rates from one stacked conformation to the other ($\tau_{\text{relx}} \approx 100\,\mu\text{s}, k \approx 10\,\text{ms}^{-1}$). Adding Mg^{2+} is known to slow down the transition rates to the ms-regime, depending on the amount of salt. However, it remains unclear if the intermediate, cross-shaped structure is a stable intermediate state or a fast decaying transition state (see sketch in Fig. 3**a**) [27–29] .

A. Slow dynamics

The long observation times in an OFA make it suitable for investigating slow dynamical processes that are not easily accessible under normal circumstances. Furthermore, larger brightness in an OFA helps to detect rapidly changing transitions more efficiently. Thus, we use a multi-parameter analysis to examine the HJ structures in buffer solution at 1 mM Mg^{2+} concentration without (Fig. 3**c**) and with OFA (Fig. 3**d**). The 2D-histograms show the signal extracted from a burst-wise analysis [30], whereby the observed translational diffusion times were $\tau_{\text{diff}} = 0.7$ ms without the OFA and $\tau_{\text{diff}} = 2$ ms using the antenna. In the absence of OFA, changes in the molecular conformation occur on a similar time scale as the burst duration, resulting in poor detection of the inter-state population. The appearance of a third population between the two stable FRET states in the MFD histogram of Fig. 3**d** evidence a transition between the two HJ conformations.

To confirm these observations, we performed Photon Distribution Analysis (PDA) in both scenarios. In PDA each burst is cut into equally sized time windows (TW) of typically 1, 2, and 3 ms length. Then, the FRET efficiency is extracted from each TW and arranged in a histogram, whereby the background has to be taken into account for an accurate description of the FRET states. Finally, a fit model including the number of the FRET states is applied to the experimental data. Figure 3**e** presents the result for measurements in bulk solution, whereby two stable states of medium FRET (MF) and high FRET (HF) as a small population between the two (magenta curve) are identified. Figure 3**f** shows the outcome of the same analysis in the presence of an OFA. In this case, it is visible that the magenta population is considerably larger than before, while the stable FRET states remain unperturbed. We verified that this experimental observation can be reproduced qualitatively in the simulated data if the diffusion time is tuned to mimic the effect of OFA (see SectionVI I).

B. Fast dynamics

The high photon collection offered by OFAs enables us to apply short binning times directly to single bursts in the experimental time trace, thus, making the dynamics visible "by the eye" without the need for post processing. As a result, we acquire access to fast transient processes that are expected to take place within 100 μs at low Mg^{2+} concentrations as known from measurements on immobilized molecules [31–34].

To showcase the improved time resolution, we performed measurements at 0.2 and 0.5 mM concentrations of Mg^{2+} in the buffer. Figure 4**a** shows a FRET trajectory assembled from 50 individual bursts that have been selected from different positions in the original time trace. Each burst's starting and ending points are indicated by dashed grey lines. The resulting FRET efficiency trajectory can then be studied using a Hidden Markov Model (HMM). Here, we have analyzed the tra-

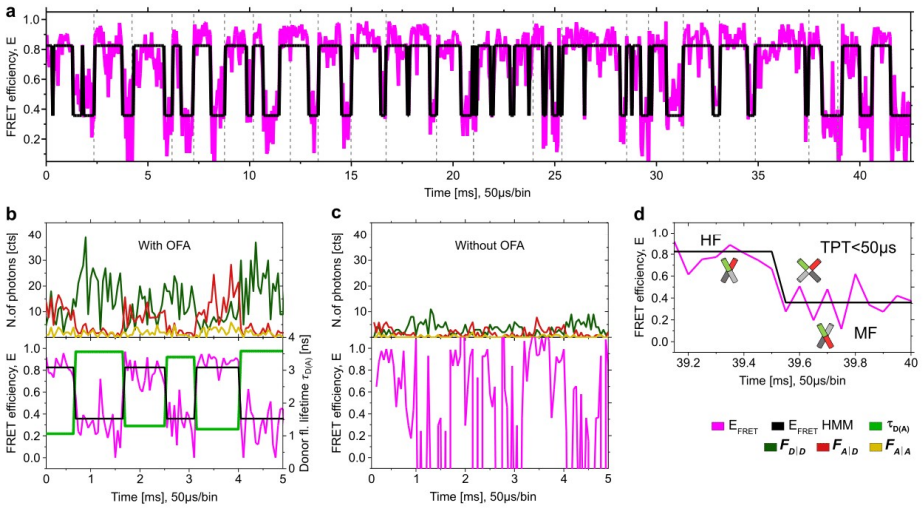


Figure 4. **HJ sm-burst trajectory. a,** FRET efficiency trace obtained by assembling individual bursts obtained from a high excitation rate experiment at $I_{\mathrm{exc}(485)} = 800\,\mu$W using a buffer containing 0.2 mM Mg^{2+}. Only bursts with a diffusion time higher than 0.8 ms that show a transition from HF to MF or vice versa were considered. The light pink line plots the data, and the black lines indicate the efficiency levels fitted using a Hidden Markov Modeling (HMM) algorithm. The dashed grey line indicates the starting/ending point of each individual burst. **b,** Close-up of the FRET efficiency histogram in an OFA. The top panel shows the fluorescence intensities. $F_{D|D}$: fluorescence from the donor under donor excitation; $F_{A|D}$: fluorescence of the acceptor under donor excitation; $F_{A|A}$: fluorescence of the acceptor under acceptor excitation. The lower panel shows the FRET efficiency trajectory (E), FRET levels obtained with the HMM analysis (E HMM) and fluorescence lifetime of the donor molecule $\tau_{D(A)}$). The lifetime was fitted for the interval where a stable FRET level was identified. **c,** Same trajectory as in (b) but with 0.75% fewer photons to mimic lack of OFA. **d** Close-up of the obtained FRET efficiency trajectory using a 50 μs time binning. The transition always occurred within a single bin resulting in a Transition Path Time (TPT) of less than 50 μs.

jectory of E with the software HaMMy from McKinney et al. [35], and a static model was implemented to fit the data. This point is justified since at binning times shorter than 100 μs, FRET levels can be mapped with several data points. It follows that dynamical averaging can be efficiently avoided at an even shorter binning time of 50 μs (for more details, see SectionVI K).

The results of the HMM analysis allow us to extract the exponentially decaying level durations of the HF and MF states directly from the trace, yielding $\tau_{\mathrm{relx}} = 0.25$ ms for the 0.2 mM concentration and $\tau_{\mathrm{relx}} = 0.31$ ms for the 0.5 mM concentration. These values closely match those obtained with an fFCS analysis (see SI section VI K) and agree with the general trend of HJ to have higher relaxation times for higher Mg^{2+} concentration [27]. We note that each burst in the composed E trajectory might end before the FRET level ends, causing a tendency towards shorter relaxation times. This effect was tested in simulations, and we found that the expected relaxation time is shortened by roughly 30% when analyzing merged bursts of single-molecule events with HMM (see SI section VI K). We remark that the E trace can be complemented with PDA analysis for accurate estimation of the relaxation rates. Furthermore, it can be beneficial for studying the transition path time of the molecule to go from one state to another.

Figure 4**b** shows a close-up of fast trajectories of the two background-corrected fluorescence channels recorded at a speed of 1.3 MHz. A visual inspection of such raw data clearly reveals the alternate modulations of the FRET levels. In addition, the donor fluorescence lifetime τ_{DA} was fitted for the interval of a stable FRET level, showing the expected anti-correlation to E. To demonstrate the advantage of the OFA, in Fig. 4**c** we randomly deleted 75% of the detected photons from the experimentally recorded data to mimic the case of no OFA (Fig. 4**c**). We find that the E levels are no longer identifiable at a binning time of 50 μs. We present a more complex comparison in the SI section VI K, where simulated traces with and without OFA are compared. The simulation shows that without OFA, the large distribution of E ($\sigma_{\mathrm{eff}} = 0.32$) prevents

the HMM algorithm from identifying reasonable FRET levels. In contrast, the FRET efficiency distribution is significantly narrower ($\sigma_{\text{eff}} = 0.14$) in the OFA, enabling the HMM algorithm to find stable FRET levels efficiently (see SI section VI K). Since we could not detect a stable third state at any point of the FRET trace, the OFA measurements let us conclude that this state involves a rather fast decaying transition with a relaxation time $\tau_{\text{relx}} << 0.05\,\text{ms}$ (see an exemplary transition in Fig. 4**e**).

IV. CONCLUSIONS

We have introduced and characterized a new optofluidic antenna, which not only enhances the optical signal obtained from emitters in a liquid, but it also gives access to longer diffusion and observation times. The fabrication and alignment of the device are inexpensive and highly fault tolerant regarding the antenna dimensions and operation spectrum. These features usher in sensitive contact-free optical measurements that give access to both faster and slower dynamics of biological entities than is available in a bulk fluidic environment. Moreover, OFAs can be combined with other platforms such as plasmonic systems [9, 36] or other methods that slow down the translational diffusion of analytes such as trapping, immobilization, or tethering mechanisms [**?** **?**]. OFA is not only broadly applicable to a variety of fluorescence measurements, but it is also compatible with other microscopy methods such as phase-contrast microscopy, dark field microscopy and iSCAT.

[1] C. Zander, M. Sauer, K. H. Drexhage, D. S. Ko, A. Schulz, J. Wolfrum, L. Brand, C. Eggeling, and C. A. M. Seidel, Applied Physics B Laser and Optics **63**, 517–523 (1996).

[2] E. Lerner, A. Barth, J. Hendrix, B. Ambrose, V. Birkedal, S. C. Blanchard, R. Börner, H. Sung Chung, T. Cordes, T. D. Craggs, A. A. Deniz, J. Diao, J. Fei, R. L. Gonzalez, I. V. Gopich, T. Ha, C. A. Hanke, G. Haran, N. S. Hatzakis, S. Hohng, S.-C. Hong, T. Hugel, A. Ingargiola, C. Joo, A. N. Kapanidis, H. D. Kim, T. Laurence, N. K. Lee, T.-H. Lee, E. A. Lemke, E. Margeat, J. Michaelis, X. Michalet, S. Myong, D. Nettels, T.-O. Peulen, E. Ploetz, Y. Razvag, N. C. Robb, B. Schuler, H. Soleimaninejad, C. Tang, R. Vafabakhsh, D. C. Lamb, C. A. Seidel, and S. Weiss, eLife **10**, e60416 (2021).

[3] B. Hellenkamp, S. Schmid, O. Doroshenko, O. Opanasyuk, R. Kühnemuth, S. Rezaei Adariani, B. Ambrose, M. Aznauryan, A. Barth, V. Birkedal, M. E. Bowen, H. Chen, T. Cordes, T. Eilert, C. Fijen, C. Gebhardt, M. Götz, G. Gouridis, E. Gratton, T. Ha, P. Hao, C. A. Hanke, A. Hartmann, J. Hendrix, L. L. Hildebrandt, V. Hirschfeld, J. Hohlbein, B. Hua, C. G. Hübner, E. Kallis, A. N. Kapanidis, J.-Y. Kim, G. Krainer, D. C. Lamb, N. K. Lee, E. A. Lemke, B. Levesque, M. Levitus, J. J. McCann, N. Naredi-Rainer, D. Nettels, T. Ngo, R. Qiu, N. C. Robb, C. Röcker, H. Sanabria, M. Schlierf, T. Schröder, B. Schuler, H. Seidel, L. Streit, J. Thurn, P. Tinnefeld, S. Tyagi, N. Vandenberk, A. M. Vera, K. R. Weninger, B. Wünsch, I. S. Yanez-Orozco, J. Michaelis, C. A. M. Seidel, T. D. Craggs, and T. Hugel, Nature Methods **15**, 669 (2018).

[4] M. Orrit, T. Ha, and V. Sandoghdar, Chemical Society Reviews **43**, 973 (2014).

[5] S. Adhikari and M. Orrit, The Journal of Chemical Physics **156**, 160903 (2022).

[6] M. Krishnan, N. Mojarad, P. Kukura, and V. Sandoghdar, Nature **467**, 692 (2010).

[7] N. Banterle and E. A. Lemke, Current Opinion in Biotechnology **39**, 105 (2016).

[8] B. Croop, C. Zhang, Y. Lim, R. M. Gelfand, and K. Y. Han, Wiley Interdisciplinary Reviews: Systems Biology and Medicine , e1445 (2019).

[9] H. Altug, S.-H. Oh, S. A. Maier, and J. Homola, Nature Nanotechnology **17**, 5 (2022).

[10] K. G. Lee, X. W. Chen, H. Eghlidi, P. Kukura, R. Lettow, A. Renn, V. Sandoghdar, and S. Götzinger, Nature Photonics **5**, 166 (2011).

[11] X.-W. Chen, S. Götzinger, and V. Sandoghdar, Optics Letters **36**, 3545 (2011).

[12] X.-L. Chu, T. J. K. Brenner, X.-W. Chen, Y. Ghosh, J. A. Hollingsworth, V. Sandoghdar, and S. Götzinger, Optica **1**, 203 (2014).

[13] C. Xiao-Liu, S. Götzinger, and V. Sandoghdar, Nature Photonics **11**, 6 (2017).

[14] K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg, Nature Methods **7**, 377 (2010).

[15] H. Qian and E. L. Elson, Appl. Opt. **30**, 1185 (1991).

[16] W. J. Rigler R., Mets Ü. and P. Kask, European Biophysics Journal **22** (1993).

[17] E. L. Elson, Methods **140-141**, 3 (2018).

[18] Z. Zhou, X. Yan, Y.-H. Lai, and R. N. Zare, The Journal of Physical Chemistry Letters **9**, 2928 (2018).

[19] H. Xiong, J. K. Lee, R. N. Zare, and W. Min, The Journal of Physical Chemistry Letters **11**, 7423 (2020).

[20] H. Xiong, J. K. Lee, R. N. Zare, and W. Min, The Journal of Physical Chemistry B **124**, 9938 (2020).

[21] M. F. Ruiz-Lopez, J. S. Francisco, M. T. C. Martins-Costa, and J. M. Anglada, Nature Reviews Chemistry **4**, 459 (2020).

[22] P. S. Cremer, A. H. Flood, B. C. Gibb, and D. L. Mobley, Nature Chemistry **10**, 8 (2018).

[23] R. J. Saykally, Nature Chemistry **5**, 82 (2013).

[24] R. Holliday, Genetical Research **5**, 282–304 (1964).

[25] E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, and C. A. Seidel, in *Methods in Enzymology*, Vol. 475 (Elsevier, 2010) pp. 455–514.

[26] E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, and C. A. Seidel, in *Methods in Enzymology*, Vol. 475 (Elsevier, 2010) pp. 455–514.
[27] S. A. McKinney, A.-C. Déclais, D. M. Lilley, and T. Ha, Nature Structural Biology **10**, 93 (2003).
[28] A. Gietl, P. Holzmeister, D. Grohmann, and P. Tinnefeld, Nucleic Acids Research **40**, e110 (2012).
[29] R. Clegg, A. Murchie, and D. Lilley, Biophysical Journal **66**, 99 (1994).
[30] S. Kalinin, A. Valeri, M. Antonik, S. Felekyan, and C. A. M. Seidel, The Journal of Physical Chemistry B **114**, 7983–7995 (2010).
[31] C. Eggeling, J. Widengren, L. Brand, J. Schaffer, S. Felekyan, and C. A. M. Seidel, The Journal of Physical Chemistry A **110**, 2979 (2006).
[32] M. Margittai, J. Widengren, E. Schweinberger, G. F. Schroder, S. Felekyan, E. Haustein, M. Konig, D. Fasshauer, H. Grubmuller, R. Jahn, and C. A. M. Seidel, Proceedings of the National Academy of Sciences **100**, 15516 (2003).
[33] H. S. Chung and W. A. Eaton, Nature **502**, 685 (2013).
[34] J.-Y. Kim and H. S. Chung, Science **368**, 1253 (2020).
[35] S. A. McKinney, C. Joo, and T. Ha, Biophysical Journal **91**, 1941–1951 (2006).
[36] H. Yuan, S. Khatua, P. Zijlstra, M. Yorulmaz, and M. Orrit, Angewandte Chemie International Edition **52**, 1217 (2013).
[37] L. A. Campos, J. Liu, X. Wang, R. Ramanathan, D. S. English, and V. Muñoz, Nature Methods **8**, 143–146 (2011).
[38] F. Mugele and J. Heikenfeld, *Electrowetting: Fundamental Principles and Practical Applications* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2018).
[39] J. R. Fries, L. Brand, C. Eggeling, M. Köllner, and C. A. M. Seidel, The Journal of Physical Chemistry A **102**, 6601–6613 (1998).
[40] S. Felekyan, S. Kalinin, H. Sanabria, A. Valeri, and C. A. M. Seidel, ChemPhysChem **13**, 1036 (2012).
[41] A. Barth, O. Opanasyuk, T.-O. Peulen, S. Felekyan, S. Kalinin, H. Sanabria, and C. A. M. Seidel, (2021), arXiv:2107.14770 [physics.chem-ph].
[42] M. Antonik, S. Felekyan, A. Gaiduk, and C. A. M. Seidel, The Journal of Physical Chemistry B **110**, 6970–6978 (2006).
[43] L. Novotny and B. Hecht., *Principles of Nano-Optics.* (Cambridge University Press, New York, United States, 2006).
[44] R. Rigler, U. Mets, J. Widengren, and P. Kask, European Biophysics Journal **22** (1993), 10.1007/bf00185777.
[45] J. Widengren, Analytical Chemistry **78**, 2039–2050 (2006).
[46] V. Kudryavtsev, M. Sikor, S. Kalinin, D. Mokranjac, C. A. M. Seidel, and D. C. Lamb, ChemPhysChem **13**, 1060–1078 (2012).
[47] J. Hendrix and D. C. Lamb, Methods in Enzymology , 205–243 (2013).
[48] S. Kalinin, T. Peulen, S. Sindbert, P. J. Rothwell, S. Berger, T. Restle, R. S. Goody, H. Gohlke, and C. A. Seidel, Nature Methods **9**, 1218–1225 (2012).

V. METHODS

Antenna preparation. The cover glass substrate model No.1.5H-170 um (tol. 5 um) was purchased from Marienfeld-Superior. Glass capillaries model **TW-100-3** where purchased from World Precision Instruments. Micropipettes were prepared in a heat-pulling procedure using a Micropipette-Puller/Sutter Instruments CO. Model-P-2000. The micropipette was dipped in dichlorodimethylsilane making its inner wall hydrophobic in order to avoid water for entering inside the pipette due to capillary forces and to control the meniscus shape of the water/air interface. The remaining dichlorodimethylsilane inside the micropipette was removed afterwards by dipping the micropipette in acetone under an ultrasonic bath for several minutes. A further step of cleaning with IPA for two minutes in ultrasonic bath is done to remove residual Acetone. The remaining IPA is then removed by heating the micro pipette for several minute at 250 Celsius degrees.

Experimental conditions. The measurement buffer for Rhodamin110 was ultrapure water, for the HJ in case of normal excitation ($I_{\mathrm{exc}(485)} = 60\mu$W) the buffer contained 10 mM Tris, 50 mM NaCl at a pH=7.5 with a varying amount of Mg^{2+}. In case of high excitation ($I_{\mathrm{exc}(485)} = 800\mu$W) additional photoprotection was used by adding 1 mM Trolox and 10 mM Cysteamine as suggested in [37].

MFD setup. Experiments were performed on a home built sm-MFD setup based on an inverted microscope with a PIE excitation scheme. Position of the antenna was calibrated using a CCD camera. Solutions were diluted to pM-concentrations to ensure single molecule events only. Further details can be found in VI F.

Fluorescence Correlation Spectroscopy. Fluorescence correlation analysis was done using unfiltered signal (free dye studies) and a fluorescence lifetime filtered signal (HJ). We applied a model including diffusion in a 3-dimensional Gaussian shaped volume with a triplet state. Brightness Q was calculated using the number of molecules in a bright state N_{bright} and fitted using an exponential saturation curve. More details can be found in VI D and VI J.

Sample preparation. Individual labeled and unlabeled single strands were ordered from IBA-Lifesciences (IBA Lifesciences GmbH, Germany). Single strands have the following sequence: α (5'-CCT AAT TAC CAG TCC AGA TTA ATC AGT ACG), β (5'-CGT ACT GAT TAA TCT CCG CAA ATG TGA ACG), γ (5'- CGT TCA CAT TTG CGG TCT TCT ATC TCC ACG), δ (5'-CGT GGA GAT AGA AGA GGA CTG GTA ATT AGG). Strand α was labeled at nucleotide 7 (Tyhmine) using Alex-a488 and strand β at 10 (Thymine) using Atto-647N, hence the hybridized

sample is called D(a)A(b). Hybridization is done using a PCR machine (Pico17, Thermo Fisher Scientific) using a concentration ratio of 1:3 labeled to unlabeled with 5nM and 15 nM, respectively. The sample was heated up to 85°C, quickly cooled down to 52°C, the temperature was held for 2 hours, and then cooled down to 4°C. The protocol leads to a typical ratio of double labeled molecules of more than 80%. Ratio was monitored using sm-MFD via stoichiometry value (see example in Figure 10). Anisotropy values were used to exclude the existence of free dye in the sample.

Monte Carlo simulations of diffusion.

Here, we assume that the excitation power is low enough that photobleaching can be neglected. The simulation starts with the analyte placed at the center of the observation volume, and $N = 10^7$ steps of diffusion are considered in a total time of 800 ms. The process was repeated 10^4 times. We used a time step of 80 ns sufficient to map the translational diffusion of the molecules. A burst is created for every entry-exit event through the observation volume (see SI for more details). The boundary surfaces of the OFA were set to direct the molecule in a random direction after each contact. The shape of the water meniscus at the air interface was approximated by an ellipse with a semi-minor axis length of 500 nm in the axial direction and a semi-major axis of 7 μm, corresponding to the dimensions of the micropipette. In the experiment, the shape of the water/air meniscus can be modified by applying pressure. This process is monitored by looking at the formation of stripes (instead of Newton rings) in the interference pattern formed due to the thin water layer. We verified that the simulation results are not sensitive to the exact shape of the interface but rather to the surface concavity. [38].

Molecular detection efficiency (MDE)

FCS and FRET measurements are performed on a confocal microscope. Fluorescence excitation is confined to a small region defined by the focused laser beam. The transmission of the emitted photons to the detector depends on the numerical aperture of the objective lens and the size of the pinhole in the image plane. The combined effect of the optical system on the signal intensity is characterized by the molecular detection efficiency (MDE). We calculate this function by modelling the excitation $I(r, z)$ with a Gaussian beam, and the image of a molecule located at r, i.e. the point-spread function PSF(r, r', z), with vectorial diffraction. The MDE is then given by

$$\mathrm{MDE}(r, z) = I(r, z) \cdot \int dr \mathrm{PSF}(r, r', z) \cdot T(r'),$$

where $T(r')$ represents the transmission through the pinhole. The parameters of the imaging system in the calculation are identical to those in the experiment. In order to assess the relative enhancement of photon detection inside the OFA for varying water layer thickness, we compare the spatial averages of the MDE.

Single molecule Multiparameter Fluorescence Detection. Single molecule bursts were identified using a burst search algorithm according to [39] using a Lee filter, a time threshold of typically 0.1 ms and a minimum of 60 photons per burst. For TW-based analysis in MFD each burst was divided into equally sized TWs of a length of 200 μs, in case of intensity based trace analysis 50 μs. Correction factors for intensity based confocal MFD were estimated using spectra of the used dyes and measured spectra of the optical components based on [3]. In short, fluorescence signal is corrected for background, crosstalk/donor leakage α, different excitation flux β and the ratio of detection efficiency and quantum yields γ. Quantum yields were estimated using the lifetime of the donor only population for the donor and the lifetime of the direct acceptor excitation for the acceptor. It follows for the fluorescence of the donor F_D after donor excitation $F_{D|D}$:

$$F_{D|D} = \gamma \cdot I^{ii}_{D|D}, \tag{1}$$

where $I^{ii}_{D|D}$ denotes the background corrected raw intensity of the donor after donor excitation. In the same way it follows for the acceptor fluorescence after acceptor excitation $F_{A|A}$ and after donor excitation $F_{A|D}$:

$$F_{A|A} = \frac{1}{\beta} \cdot I^{ii}_{A|A}, \tag{2}$$

$$F_{A|D} = I^{ii}_{A|D} - \alpha I^{ii}_{D|D} - \delta I^{ii}_{A|A}. \tag{3}$$

From these, one can estimate the FRET efficiency E and the stoichiometry S as follows:

$$E = \frac{F_{A|D}}{F_{A|D} + F_{D|D}}, \tag{4}$$

$$S = \frac{F_{D|D} + F_{A|D}}{F_{D|D} + F_{A|D}F_{A|D}}, \tag{5}$$

In this work S was used to monitor the labeling quality and hybdridization efficiency of the HJ (see 10).

FRET efficiency trajectory. FRET efficiency trajectories were assembled from single bursts taken under high excitation rate. Bursts were filtered using the already discussed sm-MFD approach based on their stoichiometry, Alex-2CDE value and differences in macrotimes to filter bleaching. Additionally a threshold of a minimum diffusion time of $t_d > 0.8$ ms was applied

since the majority of bursts were bleached after 0.2 ms due to the high excitation rate.

Hidden Markov Model. Fitting of the FRET efficiency trajectories was done using a HMM with theory and software from [35]. HMM is suitable to find FRET efficiency levels hidden in noise in a trajectory based on time binned data . It is applicable for a multi state system where each single state decays exponentially. It describes these system with transition probabilities for the sample to go from one state to the other and emission probabilities that models the distribution of FRET efficiency values with Gaussian distributions . **Simulation.** Simulated data was generated using a Brownian dynamics approach with in-house software [40] available upon request. Molecules were simulated according to the in the experiment measured FRET efficiency levels, brightness, dynamic rate and diffusion time. For the diffusion of the molecule we applied a free diffusion in a 3D Gaussian volume.

FRET-lines Static FRET lines were calculated based on [30] and [41]. In short, the static FRET efficiency line is calculated as the ratio of the donor lifetime in presence of an acceptor $\tau_{D(A)}$ and in absence of an acceptor $\tau_{D(0)}$:

$$E_{static} = 1 - \frac{\langle \tau_{D(A)} \rangle_f}{\tau_{D(0)}}. \tag{6}$$

After correction for linker dynamics molecules show FRET efficiencies/lifetimes on the static FRET line in case of no dynamic averaging in the signal. In contrast, if dynamic averaging occurs, the measured, photon weighted average fluorescence lifetime $\langle \tau_{D(A)} \rangle_f$ is biased towards longer lifetimes, shifting the population off the static line.

Photon Distribution Analysis Besides the burstwise approach also a TW approach was done. Here procedures from [42] are used. Chosen TWs size was 1, 2 and 3 ms for normal excitation rate experiments and 50 μs for high excitation rates experiments and binned in typically $N_{bin} = 81$ for FRET efficiency values E or ratios of the green and red signal S_{green}/S_{red}. Histograms are analyzed using in case of the HJ a two state model (see 3a and 4c) with dynamic rates k_{12} for HF to MF and k_{21} for MF to HF. For fitting a standard Levenberg-Marquardt algorithm was used.

12

VI. SUPPLEMENTARY INFORMATION

CONTENTS

A. Confocal volume simulation

In this section, we study the intensity profile of a focused laser beam inside the OFA. The full vectorial model of Maxwell's equations inside the dielectric structures has been solved using the "Thin lens" package from "Lumerical Inc. V.8.20.1731". As an initial input we chose a Gaussian beam polarized along the x-axis. Moreover, a lens with a NA of 1.46 or 0.4 was used to simulate the scenarios where the light is focused into a tight confocal spot and a large confocal spot, respectively. In either case, The beam was mapped using 300 plane waves, and the wavelength was set to 532 nm. Moreover, the distance of the lens from the focal plane was 1.6 μm in the case of a lens with NA 1.46, and 8 μm in the case of a lens with NA 0.4. We note that the light source was located at one focal distance in both cases. For high precision determination of the confocal volume, a mesh grid with 5 nm resolution was applied in the region corresponding to the water layer.

First, we examine the scenario where the laser beam is tightly focused inside the water layer. Figures 5 **a** and **b** show the intensity distribution corresponding to the simulation of light being focused inside the OFA using a lens with NA 1.46. The antenna geometry for this calculations is represented by cover-glass (n_{CG} =1.517), 800 nm of water (n_{water} =1.33), and air (n_{air} =1.0). For this choice of the thickness of the water layer, the electric field distribution produced by the excitation laser inside the OFA is only slightly modified with respect to the ideal Gaussian shape. Moreover, the volume defined at I_0/e^2 is 5% smaller, and the intensity maximum is 6% higher compared to the case of the confocal volume at the glass/water interface. We note that the width of the beam is defined where the intensity drops to $1/e^2$ of the maximum intensity I_0. Figures 5 **c** and **d** show the case of a focused laser beam passing through a glass/water interface (no OFA). The intensity distribution of the electric field is shown in c and **d** at the x/z ($y = 0$ nm) and y/z ($x = 0$ nm) planes, respectively. Note that the difference between the beam width at the x-z and y-z planes is due to the linear polarization of the input beam in combination with the high NA of the lens used to focus the light [43].

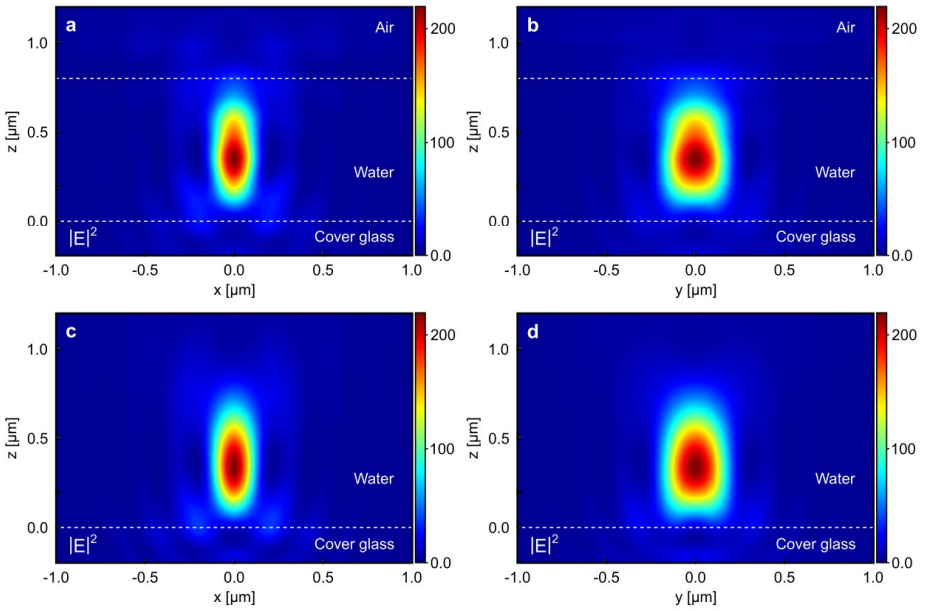


Figure 5. **Intensity profile of the excitation volume inside the OFA for NA=1.46.** Intensity distribution of the confocal spot inside the optofluidic antenna at the x/z ($y = 0$ nm) plane (**a**) and at the y/z ($x = 0$ nm) plane (**b**). The dashed white lines indicates the interfaces of the materials constituting the OFA. Intensity distribution of the confocal spot passing through the glass/water interface at the x/z ($y = 0$ nm) plane (**c**) and at the y/z ($x = 0$ nm) plane (**d**).The horizontal dashed white line indicates the cover glass/water interface.

Next, we consider the case of a large confocal spot passing through the OFA. Figures 6 **a** and **b** show the intensity distribution obtained when focusing a light beam inside the OFA using a lens with NA 0.4. The antenna geometry for this calculation uses cover-glass (n_{CG} =1.517), 500 nm of water (n_{water} =1.33), and air (n_{air} =1.0). In this scenario, the laser beam can be considered to a good approximation (to guide the reader's intuition) as a plane wave. Indeed, in the simulation we observe a modulation of the light along the z-axis produced by the interference of incident and reflected waves at the cover glass/water interface, as well as at the water/air interface. The visibility of the interference reaches 0.24 inside the water layer (see Fig. 6 **c**). Figures 6 **d** and **e** represent the intensity distribution obtained when the light passes through the water/air interface at the x/z ($y = 0$ nm) and y/z ($x = 0$ nm) planes, respectively. Finally, Fig. 6 **f** shows the intensity modulation along the z-direction in the case of focusing the light at the glass/water interface. Here, we can see that in the absence of the water/air interface, the light intensity modulation occurs only in the glass substrate, and the intensity distribution in water is that of bulk solution. In the case of a large observation volume, the main effect of the water/air interface is to restrict the analytes' diffusion. Depending on the thickness of the water layer, the observation volume can be reduced by up to tenfold (for a 500 nm water channel thickness) relative to the case of diffusion in bulk solution. Note that in the case of focusing the light with a low NA lens, the asymmetry in the x-z and y-z planes produced by the linear polarization of the input beam is no longer visible.

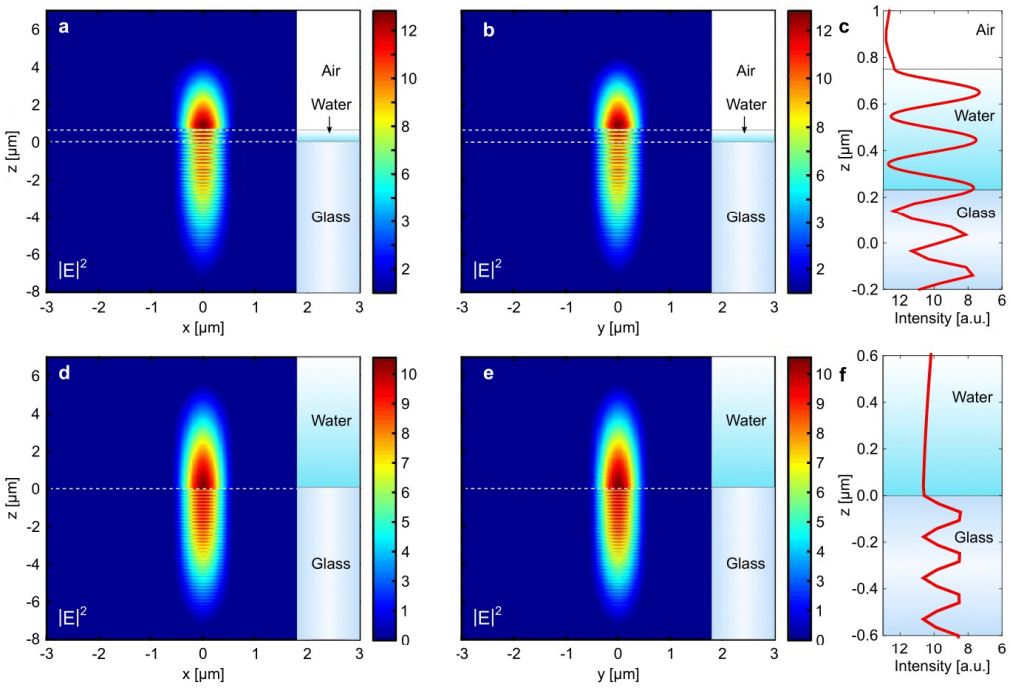


Figure 6. **Intensity profile of the excitation volume inside the OFA for NA=0.4.** Intensity distribution of the confocal spot inside the optofluidic antenna at the x/z ($y = 0$ nm) plane (**a**) and at the y/z ($x = 0$ nm) plane (**b**). The dashed white lines indicates the interfaces of the materials constituting the optofluidic antenna. **c,** Intensity distribution of the confocal spot inside the OFA along the z-direction for $x = y = 0$. Intensity distribution of the confocal spot at the glass/water interface. Figures **c** and **d** corresponds to the planes x/z ($y = 0$ nm) and y/z ($x = 0$ nm), respectively. The dashed white line indicates the glass/water interface. **f,** Intensity distribution of the confocal spot at the glass/water interface along the z-direction for $x = y = 0$.

B. Antenna diffusion: Monte-Carlo simulations

In this section, we employed Monte-Carlo simulations to investigate how the antenna geometry modifies the diffusion behavior of single molecules. In the simulations we assume that the excitation power is low enough to prevent photobleaching of the fluorophore. Figure 7 **a** shows a sketch of the structure used to simulate the antenna geometry. Here, the values for the dimensions of the micropipette are taken from the pipettes used in the experiment. Moreover, the shape of the water meniscus (water/air interface) was simulated using an elliptical function with a radii in the x and y dimension of 7 μm and a radius in the z direction of 0.5 μm. To simulate the observation volume, we defined an elliptical region in the center of the OFA with lateral radius of 0.5 μm and axial radius of 2 μm. The boundary conditions for the simulation of the diffusion of analytes inside the OFA produce a reflection in a random direction. Figure 7 **b** corresponds to the simulation in bulk solution where no boundaries are imposed. In the simulations a burst is created every time the analyte enters and exits the observation volume. Furthermore, the function '*normrnd*()' from **Matlab** was used to generate random vectors with three spatial coordinates. These vectors are then used to simulate the trajectories. For the simulation we use a time step of 80 ns and a diffusion coefficient of $D = 400\,\mu\text{m}^2/\text{s}$ for the simulations with and without the OFA. Moreover, each simulation corresponds to $N = 10^7$ steps of diffusion which is equal to 0.8 s of diffusion. Furthermore, an average of 10^4 trajectories is considered in each simulation.

The results of the simulations reveal a larger number of bursts per unit of time with the OFA compared to the case of diffusion in an open solution. Figure 7 **c** shows exemplary time traces obtained with the OFA (red trace) and in an open solution (blue trace). Here, both traces have been normalized to the maximum number of photons collected per burst. The difference and amplitude in the number of bursts in each scenario is evident. Furthermore, the dependence of the enhancement in the number of bursts obtained with the OFA as a function of the thickness of the water channel is shown in Figure 7 **d**. Here we can see that as the water channel increases, the number of bursts approaches the case of diffusion in bulk.

To scrutinize the effect posed by the water/air interface, we include in the Monte Carlo simulation a potential capable of slowing down the motion of the analytes at this interface. The idea is sketched in Fig. 7 **e**. Here, the green region represents the potential $\sigma(z)$ extending at the water/air interface in the x-y-directions. The extension of the potential in the z-direction is given by $d_\sigma = 10$ nm. Moreover, the strength of the potential is tuned by the time t_σ that the molecule stays in the region delimited by d_σ. Fig. 7 **f** shows the correlation analysis from the simulated trajectories for three different values of t_σ. We note that in the case of the dye molecule Rhodamine-110, a value of $t_\sigma = 40\,\mu$s can qualitatively reproduce the experimental observations.

Moreover, we mark that choosing the mathematical expression for the potential is not critical. For example, a step-wise function that freezes the motion of the analytes completely or a smooth function such as exponential decay of the analytes' diffusion coefficient conduce to similar results, as long as the confinement time is fixed. In this sense, the choice of the potential is degenerated since several definitions produce similar results. Finally, we note that the idea of the simulations is not to explain the physical mechanism behind the slow-down motion of the analytes but instead to give an idea of what processes can affect the motion of the analytes.

16

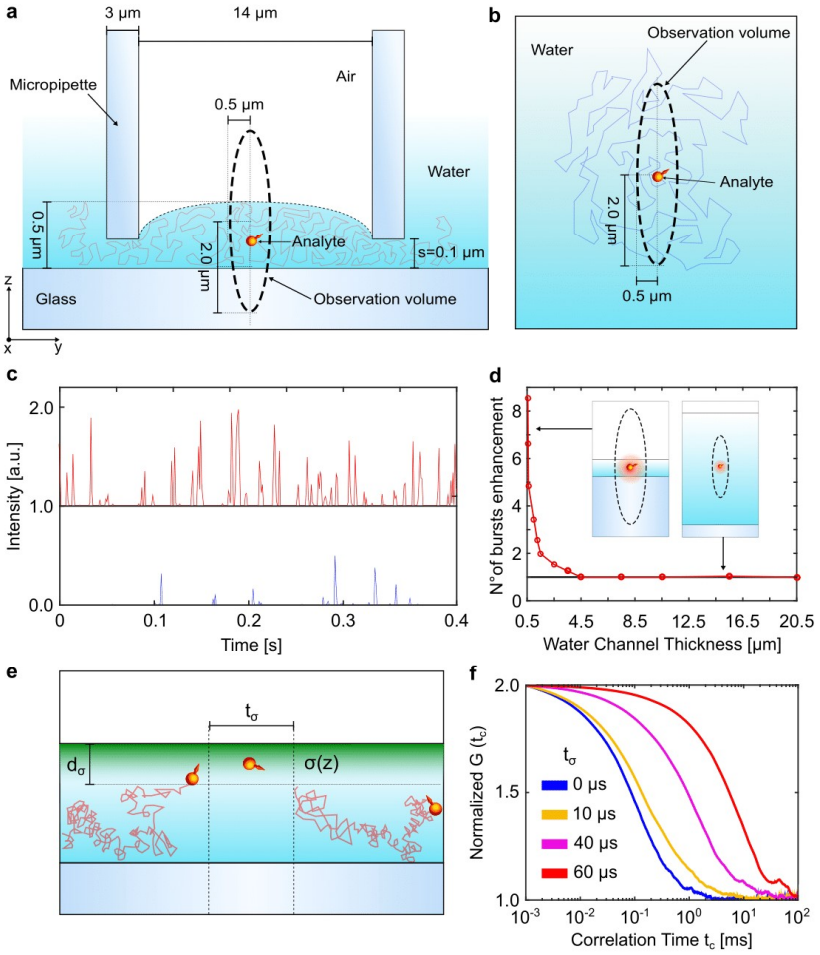


Figure 7. **Diffusion inside the OFA: Montecarlo simulation.** Sketch of the simulation where a single analyte diffuses inside the OFA (**a**) and in bulk solution (**b**). The red and blue trajectories illustrate a possible path of the analytes inside the OFA and in bulk, respectively. **c,** Exemplary time traces simulated inside the OFA (red trace) and in bulk solution (blue trace). **d,** Enhancement in the number of bursts obtained with the OFA relative to the case of diffusion in an open solution. The enhancement in the number of bursts is shown for different thicknesses of the water channel. Insert: illustration of two extreme thicknesses of the water layer. **e,** Sketch of the simulation where a potential $\sigma(z)$ (green region) is defined at the water/air interface. The potential extends along the complete interface in the x-y-directions. Moreover, the extension of the potential in the z-direction is given by $d_\sigma = 10\,\text{nm}$. The strength of the potential is modulated by the time t_σ that the molecules spend in the region defined by d_σ. **f,** Correlation analysis of the simulated time traces for different values of t_σ and a a fix value of $d_\sigma = 10\,\text{nm}$.

C. Calculation of the molecular detection efficiency (MDE)

The detection of fluorescence in a confocal microscope is characterized by the focus of the excitation beam, the point-spread function (PSF) at the image plane and the transmission function of the pinhole in the detection path [15]. The excitation beam typically has a Lorentzian-Gaussian intensity distribution

$$\begin{aligned} I(r,z) &= \frac{w_0^2 I_0}{w^2(z)} \cdot e^{-\frac{2r^2}{w^2(z)}} \\ w^2(z) &= w_0^2 + \left(\frac{\lambda}{n\pi w_0}\right)^2 z^2 \end{aligned} \tag{7}$$

where I_0 is a normalization constant. In the simulations, we set the lateral dimension of the beam waist w_0 to 1000 μm, the excitation wavelength to λ_{exc}=532 nm and the refractive index of the medium to n=1.33 in order to mimic the experimental conditions. The PSF is often approximated as that of a point-like source in a homogeneous medium. However, a more accurate description in the case of the OFA takes into account the vectorial diffraction of the electric field of a dipole in the stratified medium. We calculate the PSF for horizontally and vertically oriented dipoles using a semi-analytical approach presented by Mortensen et al. [14] for different thicknesses of the water layer ranging from 0.1 to 5 μm, a numerical aperture NA=1.46 and an emission wavelength λ_{em}=550 nm. Fig.8a shows an example of the lateral cross-sections in a 1 μm water layer. The effect of the pinhole is to limit the transmission of the signal in the image plane. The fraction of light that passes depends on the position of the image with respect to the pinhole. Thus, we define the molecular detection efficiency as

$$\mathrm{MDE}_{H,V}(r,z) = I(r,z) \cdot \int \mathrm{PSF}_{H,V}(r,r',z) \cdot T(r') \tag{8}$$

The function $T(r)$ is given by

$$T(r) = \begin{cases} 1 & r < 0.5\mu m \\ 0 & r > 0.5\mu m \end{cases} \tag{9}$$

is the transmission function of the pinhole and H, V denote horizontal and vertical dipole orientation, respectively. Fig.8b shows the axial profile along r=0 together with the excitation profile. The orientation-averaged MDE for an ensemble of molecules is then given by

$$\mathrm{MDE}(r,z) = \frac{2}{3}\mathrm{MDE}_H + \frac{1}{3}\mathrm{MDE}_V \tag{10}$$

The relative enhancement of the position-averaged MDE inside the OFA compared to an unbounded medium is shown in Fig.8c and compared with a Monte-Carlo simulation of diffusion inside through the observation volume. The parameters for $I(r,z)$ and $T(r)$ are identical in both approaches.

D. Fluorescence correlation spectroscopy

Fluorescence correlation analysis was done in different scenarios. In the case of free dye, the fluorescence signal corresponding to 2-4 molecules in the observation volume was correlated. On the other hand, In the case of experiments at the single-molecule (sm) level, between 2-5 bursts per second are registered, yielding an occupation ratio of the observation volume of roughly 1:1000. Therefore for the sm-measurements, the bursts were correlated. Moreover, to resolve dynamic events, the signal was filtered via a lifetime-based approach (see [40]). Furthermore, fitting of the correlation curves for the free dye series was done using a correlation function including the triplet state of the dye and its translation diffusion in a 3-dimensional Gaussian-shaped volume as follows [44],

$$G(t_c) = b_0 + \frac{1}{N_{bright}} \left[1 + \frac{t_c}{t_d}\right]^{-1} \left[1 + \frac{t_c}{t_d(\frac{z_0}{\omega_0})^2}\right]^{-1/2} \left[1 - |A| + |A|e^{-t_c/t_A}\right] \tag{11}$$

where b_0 denotes the offset, N the number of bright particles in the focus, t_d the diffusion time, A the amplitude for the triplet state, and t_A the time constant of the triplet state. For the free dye series, the brightness of a molecule was estimated using the number of molecules in the bright as follows,

$$N_{bright} = N \cdot (1 - A), \tag{12}$$

and therefore the brightness Q of a molecule is:

$$Q = \frac{F_{D|D}}{N} = \frac{F_{D|D}}{N_{bright}} \cdot (1 - A), \tag{13}$$

with $F_{D|D}$ the background-corrected fluorescence of the donor after donor excitation. Since increasing the excitation intensity leads to a saturation effect, we fitted the brightness as a function of the excitation intensity using a saturation curve (see Fig.2**f** in the main text),

$$Q(I_{exc}) = Q_{sat}\,(1 - exp\{I_0/I_\sigma\})\,, \tag{14}$$

where I_σ corresponds to the cross-section excitation power and is fitted globally, finally, Q_{sat} is the saturated brightness of a molecule at a maximum I_{exc}.

18

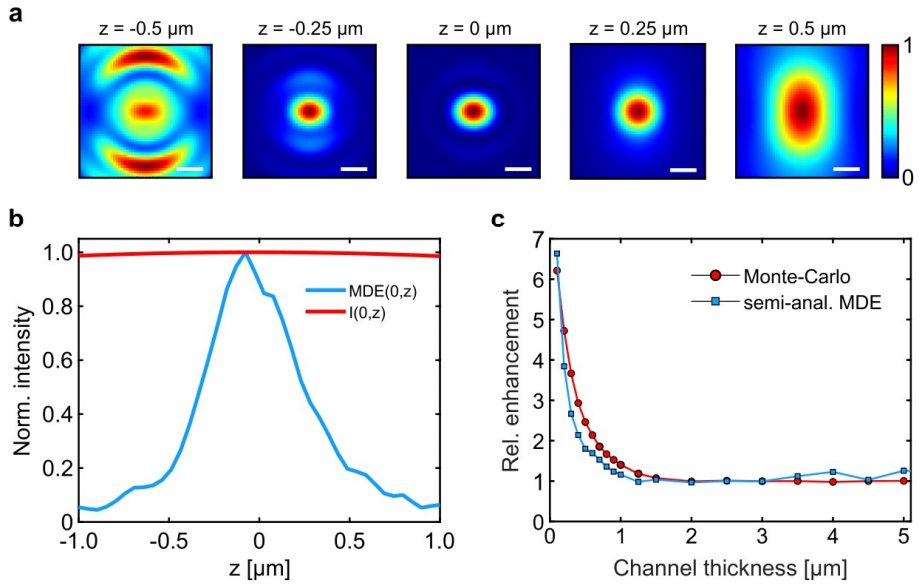


Figure 8. **Relative enhancement of photon collection in the OFA. a.** Cross-section of the PSF of a horizontally oriented dipole inside a 1 μm water layer bounded by cover-glass (n=1.5) and air (n=1). Scale bar: 200 nm. **b.** Molecular detection efficiency of the OFA (blue) and the excitation profile (red). **c.** Relative enhancement of the position-averaged MDE for small water layer thicknesses. The blue line represents a semi-analytical calculation based on vectorial diffraction, and the red line is a Monte-Carlo simulation of a diffusing molecule inside water layers with varying thickness for the same parameters of the observation volume.

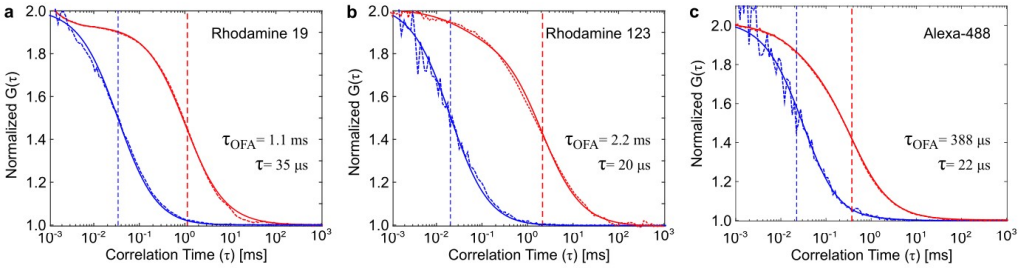


Figure 9. **Charge-dependent diffusion enhancement.** Normalized cross-correlation (dotted red, blue curves) and corresponding fit (continuous red, blue curves) obtained when measuring free dye molecules using the optofluidic antenna (red curves) and when measuring in bulk solution (blue curves). The plots from **a** to **c** correspond to Rhodamine-19, Rhodamine-123, and Alexa-488. The diffusion time in an open solution (τ) and inside the OFA (τ_{OFA}) extracted from the fits are shown as inset and also indicated with vertically dotted blue (τ) and red (τ_{OFA}) lines. All measurements were done at picomolar concentration and with 10 μW excitation power using a 532 nm (**a**) or 488 nm (**b** and **c**) laser.

E. Different charged-dye-molecules tested inside the OFA

The nature of the interaction between analytes with different charges and the water/air interface remains unknown. In the following, we test this effect in our optofluidic antenna by using dye molecules with different charges, namely, Rhodamine-19 (positively charged), Rhodamine-123 (positively charged), and Alexa-488 (negatively charged). For this experiment, the antenna geometry was fixed to a water layer thickness of 0.5 μm. Moreover, depending on the specific dye, the excitation wavelength used was a 488 nm (Toptica iBEAM SMART 488) or 532 nm (COHERENT-Sapphire 532-200). In both cases, the excitation power was set to 10 μW to avoid photodamage of the molecules. Figure 9 shows the normalized cross-correlation corresponding to the different dyes mentioned above with the OFA (red) and without using the OFA (blue). Here, positively charged molecules Rhodamine-19 and Rhodamine-123 show an enhancement in the diffusion time of 110 fold and 31 fold, respectively. On the other hand, negatively charged dye molecules like Rhodamine-110 (see Fig.2 **b** in the main text) and Alexa-488 exhibit a considerably smaller enhancement in the diffusion time of 7.5 fold and 17 fold, respectively. Finally, we mark that in this experiment, the lateral extension of the observation volume has been reduced to $w_0 = 400$ to confirm that the enhanced diffusion time is not an artifact produced by the size of the observation volume.

F. MFD setup

Single-molecule Multiparameter Fluorescence Detection (sm-MFD) experiments were performed on a home-built setup based on an Olympus IX70 inverted microscope as described in [45]. For excitation and collection of the fluorescence, we used an Olympus 100x/1.4 NA objective suitable for the OFA. For the excitation in PIE configuration a green ($\lambda_{exc} = 485$ nm) and a red ($\lambda_{exc} = 640$ nm) linearly polarized pulsed diode lasers were used (LDH-D-C 485, PicoQuant and LDH-D-C 640, Picoquant). Both lasers were operating at a repetition rate of 32 MHz [46, 47]. The fluorescence signal was filtered from the excitation light using a triple band beamsplitter 488/570/640 (AHF Analysentechnik). In the detection path, the beam is split into perpendicular and parallel polarization with a polarized beam splitter. Furthermore, the signal from the donor and acceptor molecules in each polarization channel is then separated with a dichroic beam splitter. Moreover, a further step of filtering using bandpass filters in the donor (ET535/50, AHF Analysentechnik) and acceptor channel (HQ 730/140, AHF Analysentechnik) is done to minimize the spectral overlap. Finally, each beam is directed into avalanche photodiodes (APD) for detection. Single-photon counting was done with synchronized channels (HydraHarp 400, PicoQuant, Germany) operating in Time-Tagged Time-Resolved (TTTR) mode. The data analysis was done using in-house software available upon request on the website of the Seidel group (https://www.mpc.hhu.de/software.html). Alignment of the OFA was done using an emCCD camera (DV887 ECS-BV, Andor) to position the center of the micropipette above the focus of the laser. Illumination while calibration was done using a simple torch. The objective was mounted on a

closed-loop piezo to ensure high precision and stability of the observation volume (PI, E-662).

G. HJ - sample

FRET level (state)	predicted $< R_{DA} >$	$< R_{DA} >$ with OFA	$< R_{DA} >$ without OFA
HF (stacked HJ)	36	38.4	39
MF (stacked HJ)	46	51.5	52.7
LF (planar HJ)	54	-	-

Table I. Comparison of the calculated distances using FPS to the in PDA measured distances for the measurement with 1 mM Mg^{2+} from III A

A label scheme to resolve the two stacked conformations of the HJ (HF and MF) and the potentially planar conformations (LF) was chosen based on FRET positioning screening using in-house software [48]. The measured distances based on Photon Distribution Analysis (PDA) (see I) for the HF and MF state were slightly higher than expected. Moreover, the labeling quality was monitored using stoichiometry and anisotropy. As one can see in Fig.10**a** most of the molecules were double labeled and only minor fractions of donor-and acceptor only molecules appeared. The anisotropy decay was fitted using a two component Perrin equation,

$$r_D(\tau_{D(A)}) = r_0 \left[\frac{X}{1 + \frac{\tau_{D(A)}}{\rho_1}} - \frac{1-X}{1 + \frac{\tau_{D(A)}}{\rho_2}} \right], \tag{15}$$

where X is a fraction, ρ_1 the rotation of the dye and ρ_2 the rotation of the HJ. As expected, the anisotropy decay is mostly (X=0.87) described by the rotation of the dye.

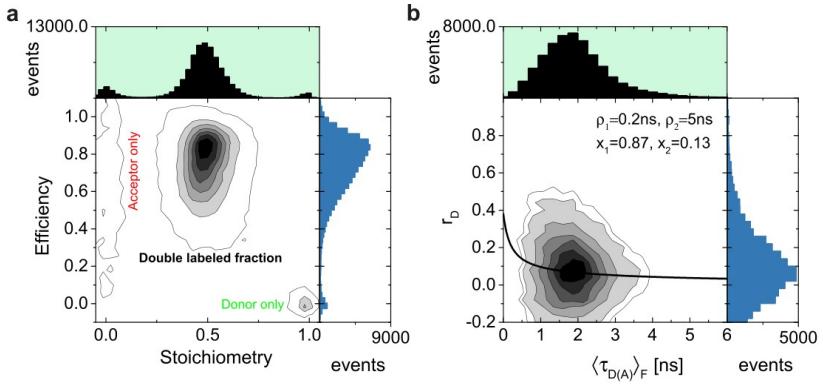


Figure 10. **Monitoring the hybridization and labeling quality of the HJ. a,** The MFD-histogram of the stoichiometry and the FRET efficiency shows three populations, two minor, this is, the acceptor only and donor only, and one major population corresponding to the double labeled HJ. **b,** Lifetime-anisotropy diagram with an overlapped two component Perrin equation (solid black line). Here, one main population of the double labeled HJ is obtained.

H. HJ-sm-MFD

Figure 11 **a** and **b** show the comparison of the HJ dynamic in an open liquid (purple) and inside the optofluidic antenna (yellow) for a buffer solution without the addition of Mg^{2+} (**a**) and for a buffer solution at 1 mM Mg^{2+}

concentration. We observe no change in the dynamic behavior of the HJ for the case of diffusion inside the OFA compared to the case of the bulk solution.

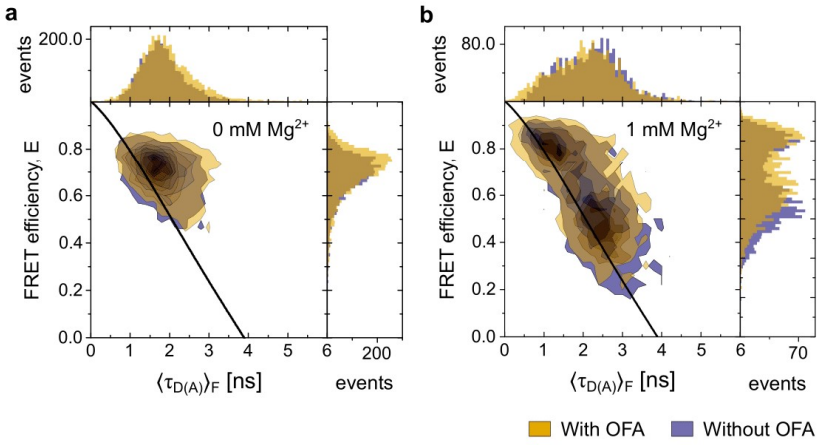


Figure 11. **Holliday junction single molecule multiparameter fluorescence detection (sm-MFD): Comparison between measurements with and without using the OFA. a,** Two dimensional histogram of the donor fluorescence lifetime in presence of an acceptor $\langle \tau_{D(A)} \rangle_F$ and the FRET efficiency E. The black line (static FRET line) describes the relation of intensity and donor lifetime based equal FRET efficiencies and bursts from molecules with a single conformation state are distributed on this line. The molecules show a fast dynamic between their conformational states and therefore a high averaging of the FRET efficiency resulting in only one visible population. The plot in purple corresponds to the case of an open solution without OFA. **b,** Adding 1 mM Mg^{2+} slows down the dynamical behavior and two populations. This effect become visible in both cases.

I. HJ - PDA

In order to test the effect of the prolonged diffusion time τ_d on the PDA, simulations similar to [30] were performed. We generated data for a two-state system with FRET efficiencies equal to the ones measured and listed in the table I. We note that in contrast to the real data, in the simulations, we assumed that all molecules show dynamical averaging. Figure 12 shows the resulting distribution of burst duration with and without the OFA for the experimental data (**a**) and for the simulated data (**b**). The simulation can reproduce the distributions for the measurements qualitatively with and without the OFA. We note that in the simulations, the diffusion time without the OFA is set to $\tau_d = 0.7$ ms and $\tau_d = 2$ ms in the case of the OFA. Figure 12**c** shows the ratio of dynamic molecules detected with the PDA as a function of the diffusion time. The percentage of active molecules detected with PDA is lower for short diffusion times than for longer detection times in a global analysis of time windows (TWs) with the size 1, 2 and 3 ms. Moreover, the simulation results show that a difference of 3 in the diffusion time results in a 10 % improvement in detecting dynamic molecules. Furthermore, in the case of molecules with different brightnesses, a difference in the number of dynamic molecules is noticeable for short diffusion times (see Fig. 12**d**). The PDA without and with OFA is shown in Figure 12**e** and **f**. Here, it is evident that molecules with longer diffusion time exhibit a higher fraction of dynamic molecules relative to the stable FRET states. Thus, the simulations indicate that the enhanced detection of dynamic molecules is mainly due to the longer observation times, which can explain the observations made in 3 qualitatively.

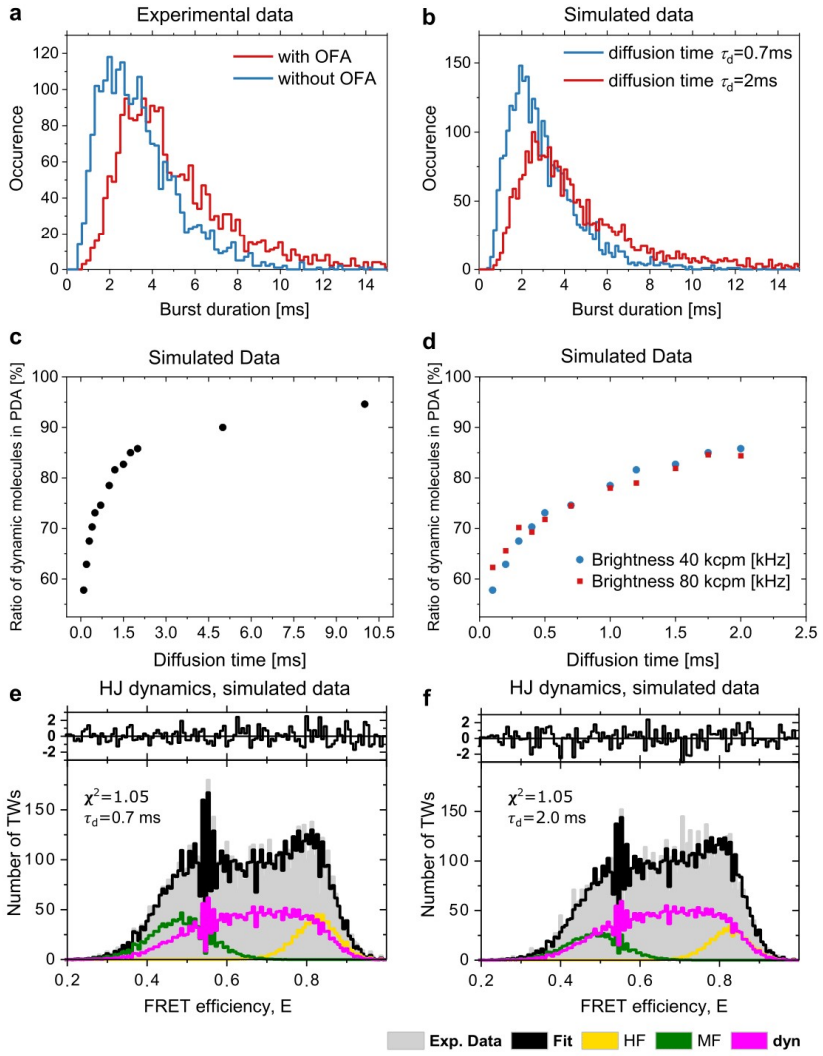


Figure 12. **Simulation of a PDA analysis using a two-state system with different diffusion times. a,** Shows the burst duration distribution obtained from the experimental data for the measurement with and without OFA. **b,** Simulated distribution using a diffusion time of 0.7 ms for the case without OFA and 2 ms for measuring with OFA. **c,** Ratio of dynamic molecules found in the PDA analysis from a simulation using only dynamic molecules in dependency of the diffusion time. **d,** Same as **c** but with molecules having different brightness. PDA analysis of the simulated two-state system at fast (**e**) and long (**f**) diffusion times. The distribution of the static HF state is shown in yellow, the distribution of the static MF state is shown in dark green, and the dynamic distribution is shown in magenta. The analysis was done using a global fit with TWs of 1, 2 and 3 ms.

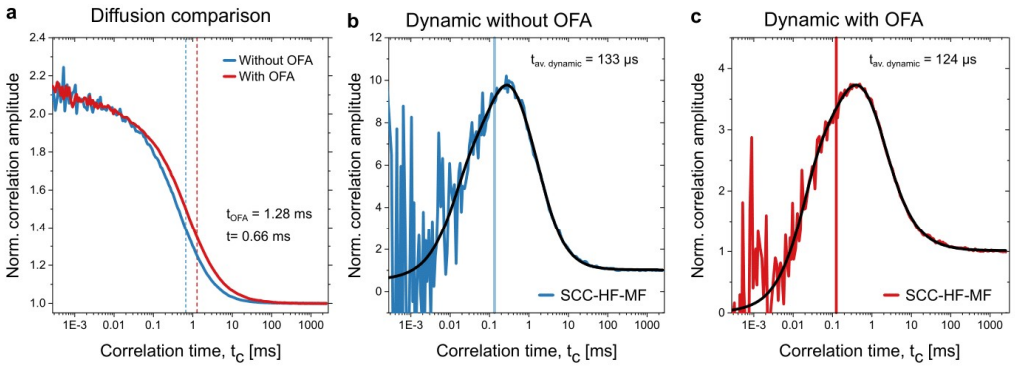


Figure 13. **HJ dynamics analysis using species-filtered Fluorescence Correlation Spectroscopy. a,** Lifetime filtered, medium FRET (MF) species auto-correlation function of the Holliday junction D(a)A(b) without Mg^{2+} with (red curve) and without (blue curve) the OFA. The vertical blue and red dashed lines indicates the diffusion time obtained from the corresponding fit, also shown numerically as inset. Figures **b** and **c** corresponds to the High FRET (HF) and medium FRET (MF) species cross correlation of the HJ without (blue curve) and with (red curve) antenna. The vertical blue and red lines indicates the average relaxation time also show as inset.

J. HJ - filtered FCS

Another approach to resolving fast dynamics between conformations is the filtered Fluorescence Correlation Spectroscopy (fFCS) based on [40]. Here, FRET species are separated using a lifetime-based filter. Selecting the HF and MF species of the HJ via burst selection leads to a distribution of short lifetimes for the HF state and longer lifetimes for the MF state. Moreover, applying these filters to the correlation analysis of all selected FRET bursts (a mixture of MF and HF) leads to a correlation curve between the FRET states. Furthermore, the correlation analysis gives information about the relaxation time between the HF and MF states in the form of an anti-bunching term in the correlation curve.

To analyze the experimental data, we consider a global-fit approach using species-auto-correlation curves (SAC) and species-cross-correlation curves (SCC). The fit of the SAC was done using equation 11 and for the fitting of the SCC we used the following equation,

$$G(t_c) = G(t_c) = b_0 + \frac{1}{N_{bright}} \left[1 + \frac{t_c}{t_d}\right]^{-1} \left[1 + \frac{t_c}{t_d(\frac{z_0}{\omega_0})^2}\right]^{-1/2} \left[1 - |A| + |A| \cdot e^{-t_c/t_A}\right] \times \tag{16}$$

$$\left[1 - |AB| \cdot \left(|B|e^{-t_c/t_B} + |C|e^{-t_c/t_C} + (1 - |B| - |C|) \cdot e^{-t_c/t_D}\right)\right], \tag{17}$$

where AB corresponds to the total amplitude of all antibunching terms and B, C, and $(1 - B - C)$ are the individual amplitudes of the different species. Analysis of the SAC shows that the diffusion time τ_d is almost three times higher using the OFA (see Fig. 13**a**). On the other hand, the average dynamic relaxation time $t_{av,dynamic} \approx 100\mu s$ show no significant change when measuring the HJ inside the OFA compared to an extended solution (see Fig. 13**b,c**). The combination of the MFD and fFCS analysis ultimately demonstrates that the use of the OFA affects neither the conformational states of the HJ nor the internal dynamic exchange rate between them.

K. HJ - FRET efficiency trajectories

The aim of the measurements using the OFA is to resolve the FRET levels of the molecule on a timescale where the states of the HJ are stable. In this regime, dynamical averaging can be avoided; thus, we can follow the switching between the FRET levels and, therefore, the transition of the HJ to another state in so-called

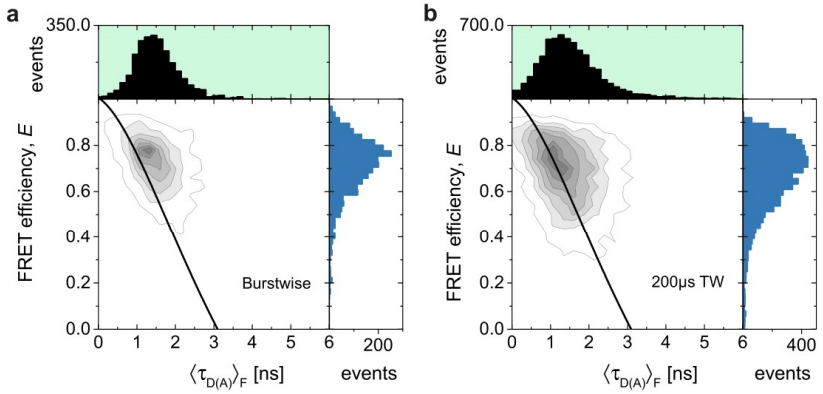


Figure 14. **Comparison of burstwise and time-window based analysis of the HJ in MFD. a,** Burstwise analysis of the sm-MFD measurement of the HJ labeled with Alexa488/Atto647N for a buffer without addition of Mg^{2+}. 2-D histogram of the lifetime of the donor in presence of an acceptor $\langle\tau_{D(A)}\rangle_F$ and the FRET efficiency E. The static FRET line was estimated using $E(\langle\tau_{D(A)}\rangle_F) = 1-((0.0126\langle\tau_{D(A)}\rangle_F^4 - 0.148\langle\tau_{D(A)}\rangle_F^3 + 0.5521\langle\tau_{D(A)}\rangle_F^2 + 0.3416\langle\tau_{D(A)}\rangle_F - 0.0174)/3.1)$. The FRET population is clearly shifted off the static FRET line towards a longer lifetime. **b,** TW analysis of the same bursts selected in the burtwise approach (see **a**). Here, every burst is cutted into equally 200 μs long TW. The 1D projections show a broadening of the distribution and the TW-bursts no longer are shifted away from the static FRET line. The excitation power used in the experiment is $I_{exc,485} = 800\,\mu$W.

"real-time". However, the poor photon detection of standard optical configurations often implies long integration times when using MFD. For example, in burstwise based analysis, the signal is integrated over the whole burst duration. Thus, when the exchange rate between molecule's conformations occurs faster than the diffusion time, the burstwise analysis averages out the FRET efficiencies levels into one population. An approach that is not limited to the burst duration is PDA, which is based on equally sized time windows with a global analysis over different TWs, typically 1, 2 and 3 ms. However, if enough photons are collected, one can also look at the MFD analysis using time windows. Figure 14 shows a comparison between a burstwise based analysis and a time window based analysis for measurements with the OFA. It is visible that the burstwise analysis shows a FRET population mainly off the static FRET line. In contrast, the FRET population is mainly on the static FRET line in the time window-based analysis. This means that the dynamical averaging can be avoided at binning times equal to or lower than 200 μs.

We took this idea further to study FRET efficiency trajectories. Figure 15 shows two such measurements at different Mg^{2+}-concentrations. Figures **a** and **e** show FRET efficiency trajectory assembled out of 70 individual bursts using 0.2 and 0.5 mM Mg^{2+}, respectively. Moreover, the FRET efficiency trajectories were analyzed with an HMM where only two states were sufficient to fit the experimental model appropriately. Figure 15**b** shows a PDA with of TWs of the same size as the time binning (50 μs). It can be seen that the HF state is most occupied, decreasing in the case of adding more Mg^{2+} as shown in Fig. 15**f**. The exponential decay of the level duration of the HF and MF states shows a longer relaxation time when adding more Mg^{2+}, as expected. This can be seen by comparing Fig. 15**c** and **g** or **d** and **h**. These observations indicate that with the OFA, it is possible to resolve FRET efficiency levels and their trajectory from single-molecule bursts originating from freely diffusing molecules.

To further test and validate this approach data was generated for high excitation rates by setting the brightnesses of the molecules to $Q = 500$kcpm (kilo counts per molecule) in the case with OFA (see Fig. 16**a**), and $Q = 150$kcpm without OFA (**b**). The relaxation times of the levels were obtained experimentally (see Fig. 15). As one can see, the HMM algorithm fails to find a reasonable amount of HF states in case of no OFA due to a lack of photons. Zoom in the simulated signal obtained with the OFA is shown in (**c**). Here, it is evident that the FRET efficiency levels are identifiable "by the eye"; thus easily fitted with the HMM algorithm, whereas, without the OFA, the FRET efficiency is too noisy for the HMM algorithm to converge (see Fig.16**d**). The analysis using HMM finds a distribution width of $\sigma_{Efficiency} = 0.14$ in case of using the OFA and $\sigma_{Efficiency} = 0.32$ without the OFA.

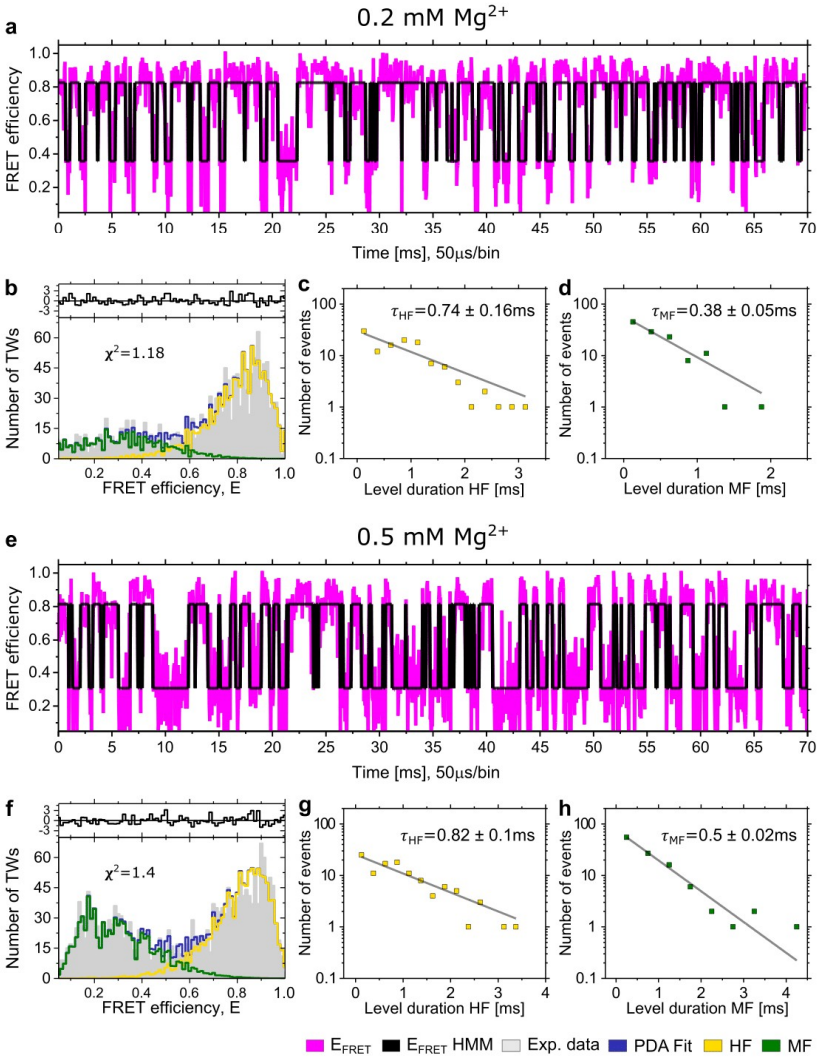


Figure 15. **Holliday Junction FRET Efficiency trajectories analysis at different Mg^{2+} concentrations. a,** FRET efficiency trajectory consisting of single bursts measured in a buffer at 0.2 mM Mg^{2+} concentration. The efficiency levels (black curve) were fitted using a HMM algorithm. **b,** histogram of the trajectory shown in **a** which as been divided into Time Windows with a binning size of 50 μs. The resulting histogram is then fitted using PDA. The dark-green distribution shows the MF-state of the HJ and the yellow distribution shows the HF-state of the HJ. The figures **c** and **d** show the distribution of the level durations of the HF-state (**c**) and the MF-state (**d**) in a semi logarithmic plot. The distribution was fitted using an exponential decay without an offset (gray curve). **e,** FRET efficiency trajectory consisting of single bursts measured in a buffer at 0.5 mM Mg^{2+} concentration. **f,** Resulting histogram from **e** with PDA using Time Windows with 50 μs binning time. The figures **g** and **h** show the distribution of the level durations of the HF-state (**g**) and the MF-state (**h**) in a semi logarithmic plot.

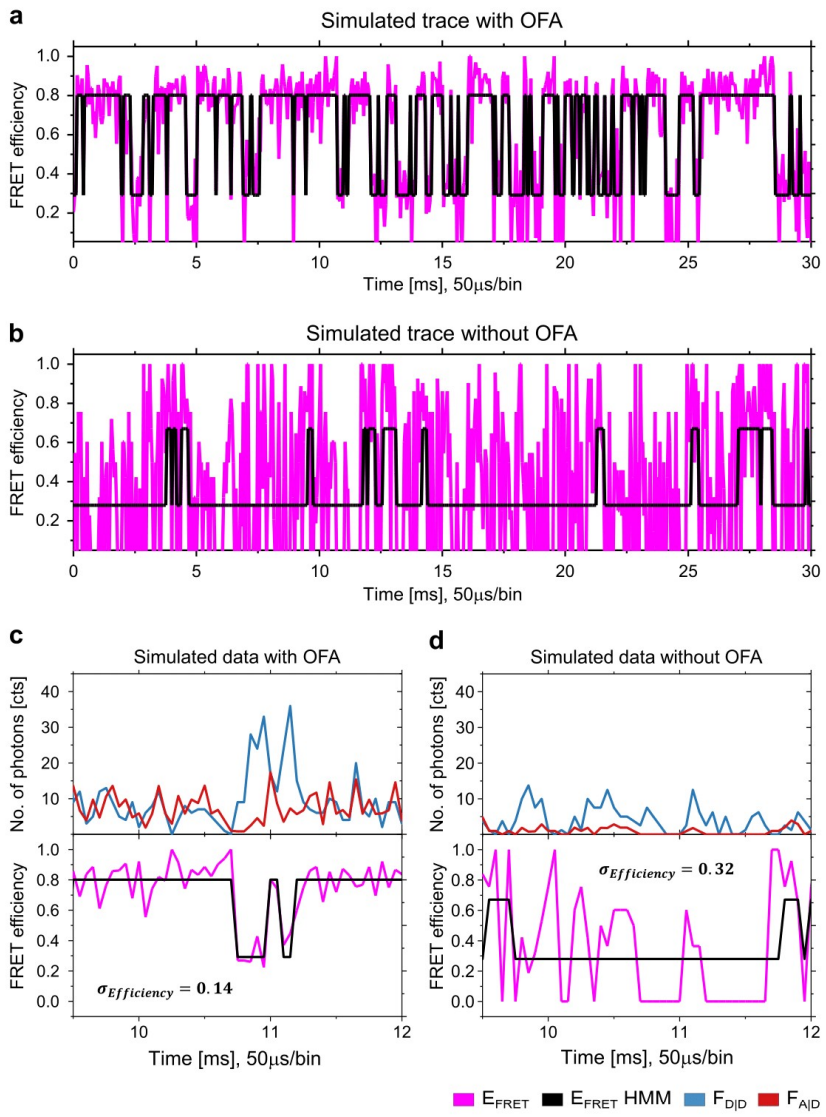


Figure 16. **Simulated FRET Efficiency trajectories with and without OFA. a,** FRET efficiency trajectory simulated considering the rates obtained from measurements at 0.2 mM Mg^{2+} concentration. In the simulation we use a two state model for the HF and MF states of the HJ. Sm-bursts were filtered for a minimum burst duration of 0.8 ms and merged to a single trajectory lasting around 70 ms. FRET levels were fitted using a HMM algorithm. **b,** The same simulation as in **a** was repeated using molecules with 1/4 of the brightness to mimic the case of no OFA. FRET efficiency was fitted again using a HMM algorithm. Figures **c** and **d** show a Zoom in the FRET efficiency trajectory with the OFA (**c**) and without the OFA (**d**). The top panel in both plots show the FRET efficiency trajectory of the fluorescence for the donor under donor excitation $F_{D|D}$ and the fluorescence of the acceptor under donor excitation $F_{A|D}$. The resulting $\sigma_{Efficiency}$ obtained from the HMM-fit is indicated on as insert in each plot.

7.3 Intramolecular dynamics and function of the human Guanylate Binding Protein 1 are controlled by farnesylation

Abstract

The human Guanylate Binding Protein 1 (hGBP1) belongs to the dynamin superfamily of large GTPases and it plays a key role in human cells´ innate immune response to bacterial, viral and protozoan attacks. In addition, antitumor activities of hGBP1 are reported. hGBP1 is only synthesized in larger amount after stimulation of the cell by interferons, otherwise the concentration is very low. As for many other GTPases an important feature for the proper function of the protein is the posttranslational, covalent attachment of an isoprenyl moiety to the protein. In the case of hGBP1 it is a farnesyl group which is linked to the C-terminus of hGBP1 and which is reported to serve as an anchor for membrane binding on the one hand and for homotypic oligomerization on the other hand. To resolve the influence of farnesylation on hGBP1 we apply sm-FRET studies of non-farnesylated and farnesylated hGBP1 and investigate its structural changes in monomeric, dimeric and oligomeric states. We find an equilibrium of different and dynamical active monomeric states for non-farnesylated hGBP1 which are not observed for the monomeric farnesylated hGBP1, where hGBP1 is in a single locked state. This changes when farnesylated hGBP1 bound to nucleotide substrate is investigated. We observe dynamic structural changes in the C-terminal part of hGBP1 which are enabled only after dimerization of the farnesylated protein. This dynamical behavior is similar as observed for monomeric non-farnesylated hGBP1. Ultimately, we trace the final step of farnesylated hGBP1 in free solution where it undergoes complete opening to form its oligomeric state. This shows that hGBP1 works as a molecular machine that requires both substrate and binding partners to achieve its function. We conclude that the C-terminal farnesyl group does not merely serve as a membrane anchor but rather it controls the structural dynamics of the protein – in the end fine-tuned by the actual concentration level of hGBP1 as only dimerization leads to the release of the farnesyl group and in turn to oligomerization and membrane binding, respectively.

Introduction

Isoprenylation is a key posttranslational modification of proteins at their C-terminal cysteine residue with either a C15 (farnesyl) or C20 (geranylgeranyl) isoprenoid lipid catalysed by Farnesyltransferase (FTase) and Geranylgeranyltransferase I (GGTase I), respectively [1]. This modification is intended to allow proteins to bind to membranes, but has also been proven to be crucial for localization, activity and protein-protein interactions [2]. One prominent example for the importance of isoprenylation is the farnesylation of the small GTPase Ras, which is a target protein in cancer therapies [3]. An interferon-stimulated, human immune protein which exhibits antitumor activities, but also responses to bacterial, viral and protozoan attacks is the human Guanylate Binding Protein 1 (hGBP1) belonging to the dynamin superfamily of large GTPases [4-7].

The defense mechanism involving GBPs is often connected to several forms of initiated cell death like pyroptosis and apoptosis [8, 9]. Human GBP1, -2 and -5 can be isoprenylated through their C-terminal CaaX-box motif, which allows for membrane binding [10]. In the case of hGBP1, where membrane association and the formation of dimers triggered by GTP binding and hydrolysis is best characterized, the added farnesyl moiety also allows for polymerization [11]. The X-ray structure of hGBP1 can be separated into several domains, which are typical for dynamin like proteins. It has a large globular GTPase domain (LGD) followed by a purely helical middle domain (MD) consisting out of two helix bundles (α7-α11) and a GTPase effector domain (GED) having two helices (α12-α13), which fold back to contact the LGD [12]. With the help of x-ray structure analysis the farnesyl moiety was shown to be bound to a hydrophobic pocket [13]. High counts of positively charged side chains on the LGD and negatively charged side chains on the GED are the foundation for the salt bridge contacts established between the GED and the LGD, resulting in a closed conformation in nucleotide free form [12]. The farnesyl moiety was reported to strengthen this closed conformation leading to changed hydrolysis behavior and dimer formation kinetics [14], but the exact impact of the farnesylation on the dynamics and structure of the protein remains unknown.

To unveil the direct effect farnesylation has on the protein, time resolved single molecule Förster Resonance Energy Transfer (smFRET) studies is the method of choice. Through this approach, it is possible to measure the protein in solution while continuously monitoring the dynamics displayed by hGBP1. Using smFRET studies to reveal conformational dynamics in proteins has been recently shown to be reliable and accurate [15]. By supplying specifically labelled samples with thus gaining quantitative distance information between fluorophores, it is possible to simulate the movement of hGBP1 using a FRET network consisting of different, intramolecular FRET pairs. The information quality of these experiments is further enhanced by using a Multiparameter Fluorescence Detection (MFD) [16] approach in combination with Pulsed Interleaved Excitation (PIE) [17].

With these sophisticated FRET methods, we show the direct influence of farnesylation on hGBP1 dynamics and structure. This further allows us to make statements about the protein's pathway during activity in both its non-farnesylated and farnesylated state in areas which were poorly understood before.

Results

Monomeric non-farnesylated hGBP1 exhibits extensive dynamics

As the goal of this work was to determine the impact of the farnesylation on the behavior of the protein as a whole, it was paramount to have a detailed understanding of the monomeric hGBP1 in terms of its structure and conformational dynamics. To do so, we investigated the non-farnesylated hGBP1 in high detail using an extended FRET labeling network (see **Figure 39a**) in combination with a label network used in a previous study (see **Figure 39b**). The low concentrations of protein in single molecule experiments (20 pM) guarantee that only monomeric protein is measured with no interference of potential dimer partners. Consecutively, we compared our findings with the farnesylated hGBP1. Here, a FRET labeling network was established to observe the changes after farnesylation (see **Figure 39c and d**). Label positions were chosen in such a way that the movements of all domains of hGBP1 against each other can be addressed, including α13 against α12, and that crucial results of the non-farnesylated hGBP1 studies can be directly compared side-to-side.

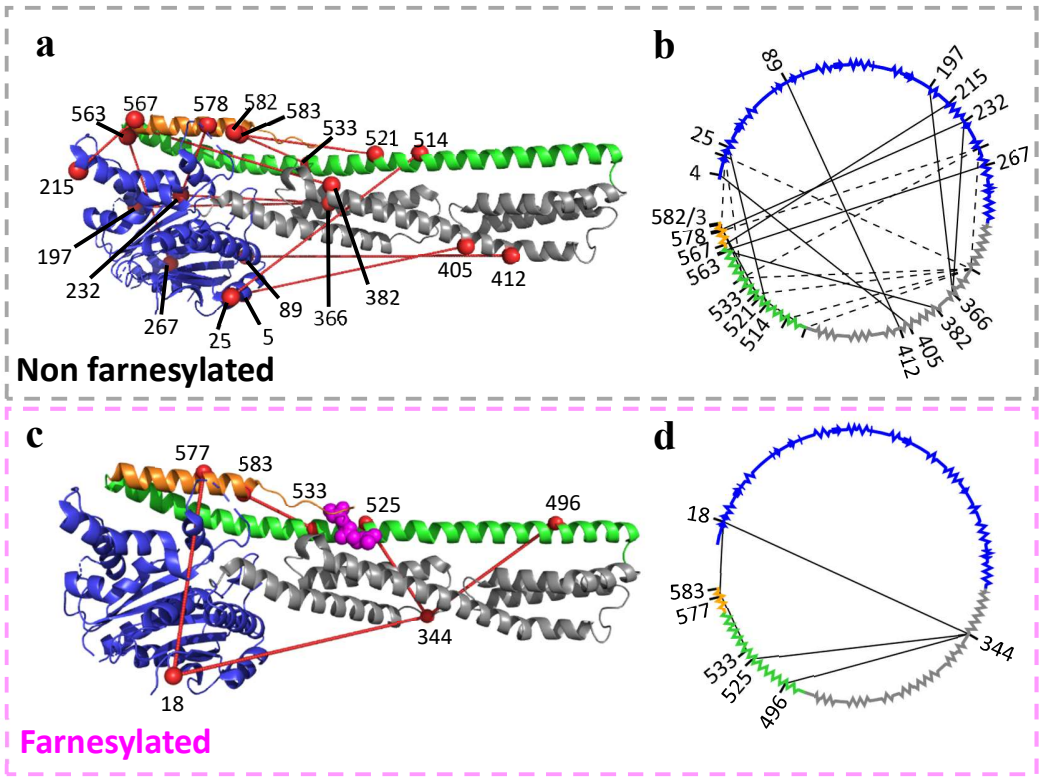


Figure 39: Labeling network of non-farnesylated and farnesylated hGBP1. a Crystal structure of the hGBP1 (PDB:6k1z) without showing the farnesyl moiety, representing the labeling network for the non-farnesylated hGBP1. LGD is shown in blue, MD in grey, α12-helix in green and α13-helix in orange. The C_{α}s of labeled amino acid positions are shown as red spheres, with the connecting lines representing a FRET pair. **b** Circular labelling scheme based on the sequence of the protein with the same FRET pairs. In addition to the FRET pairs shown in a, the FRET pairs of the previous study have been included as dashed lines. **c** Crystal structure (PDB:6k1z) with the labeling network for the farnesylated hGBP1 and **d** its corresponding circular plot.

To refine the previously determined structures of monomeric non-farnesylated hGBP1 in solution [18], the existing labeling network was extended using an algorithm described in [19]. In short, the algorithm finds optimal FRET pairs to resolve the structure of a given molecule using prior knowledge like a crystal structure, or, in this case, already derived structures from multiple experiments based on a hybrid approach. Optimal FRET pairs should find suspected new states of the molecule, be in approximate distance of the Förster radius, and not be redundant in respect to each other. Based on this algorithm, 11 new FRET pairs were identified (see **Supplementary Figure 13**), labeled (for full name and labeling efficiencies see **Supplementary Table 7**) and measured in sm-MFD and ensemble based Time Correlated Single Photon Counting (TCSPC) (see **Supplementary Figure 14** and **15**). To accommodate previous findings, data was analyzed using a two-state model where one state has a higher population in respect to the other one, referred to as M_1 and M_2, respectively. As it can be seen in a later stage of this study, the two state model was only applied to specific FRET samples, whereas others showed a more complex behavior. Measured distances and calculated errors based on the ensemble TCSPC measurements are shown in **Supplementary Table 8** and **9**. Distances and their errors were used to extend existing data which included N_{FRET} = 12 FRET pairs [18], resulting in N_{FRET} = 23 in total. The extended FRET data was additionally combined with previously taken data obtained from other techniques which were neutron spin echo, X-ray scattering and EPR spectroscopy. This extended data set was used in the same way like in the previously established global analysis [18], which works as follows: Since the data shows significant deviations to the crystal structure, new structures were generated using a combination of molecular dynamic simulations and a consecutive rigid body docking. After correcting stereochemistry using NMSim [20], resulting models were clustered and used for short MD-simulations. With another clustering step, models were discriminated using a meta-analysis by fusing the data based on Fisher's method. Finally, models were discriminated using a *p*-value of 0.68. This process was repeated using the extended data set. The resulting ensemble of structures consisted mainly of refined M_1 and M_2 states (see **Figure 40** top), compared to the original data set. As a result, the ensemble of structures that were not discriminated is significantly narrower for the M_1 state and for the M_2 state (see **Supplementary Figure 16**). The average RMSD for the M_1 ensemble is 6.3 Å (11.2 Å with original data set) and 8.2 Å for the M_2 ensemble (14.5 Å with original data set). We denote that the predicted RMSD based on the optimal FRET pairs algorithm (see **Supplementary Figure 17** based on [19]) is smaller. However, since both states M_1 and M_2 each represent an ensemble of structures which undergo dynamic equilibration, a finding of two "true" structures is not likely. While the ensemble became smaller, the overall conclusion of having two different monomeric species, with the main difference of a switched α12-helix, remained the same. Besides these two main structures being in a conformational dynamic exchange we also found a small fraction (around 15 %, see

Supplementary Table 10) of a static structure. It is similar to the M_1 state with the α13-helix attached to the α12-helix and being completely static within our observation time (see **Figure 40** left structural ensemble). This structure only became resolvable due to new labeling positions monitoring the behavior of the α13-helix in respect to the α12-helix in greater detail (samples E533C/M583C and E521C/K582C). We will call it $M_{1,\alpha13\ attached}$, whereas the other two main ensembles will be referred to as $M_{x,\alpha13\ near}$. Both names refer to the position of the α13-helix in respect to the α12-helix. On top of this, two additional sub-ensembles for $M_{x,\alpha13\ near}$ were found where the α13-helix is further away in respect to the α12-helix, as will be discussed later. The $M_{1,\alpha13\ attached}$, $M_{1,\alpha13\ far}$ and $M_{2,\alpha13\ far}$ sub-ensembles were selected via a screen of the $M_{x,\alpha13\ near}$ ensembles based on the additional resolved high FRET distance and a χ^2-threshold (see **Supplementary Figure 18**, **19,Supplementary Figure 20** and**Supplementary Figure 21**).

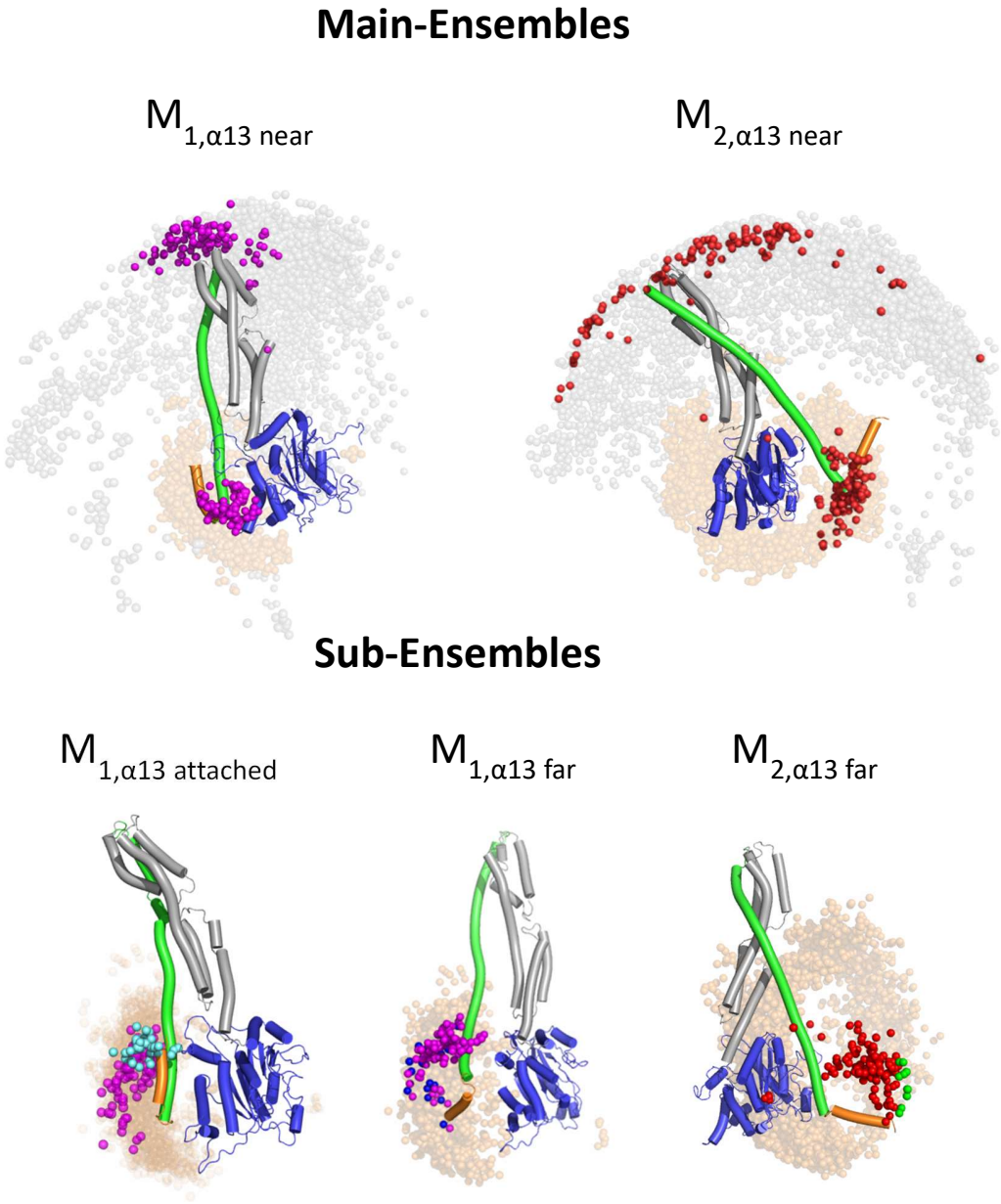


Figure 40: Structural refinement of non-farnesylated hGBP1. Structural ensemble of hGBP1 monomeric states $M_{1,\alpha13\ near}$ and $M_{2,\alpha13\ near}$ with the resolved sub-ensembles $M_{1,\alpha13\ attached}$, $M_{1,\alpha13\ far}$ and $M_{1,\alpha13\ far}$. $M_{1,\alpha13\ near}$ and $M_{2,\alpha13\ near}$ were resolved using a Meta-Analysis of EPR, SAXS and FRET including N_{FRET} = 23. Best matching structure to the experiments is shown as a cartoon. Domains are highlighted by color (blue: LGD, gray: MD, green: α12, orange: α13). Orange and grey spheres represent the whole ensemble marking the position of the last amino acid of the MD (gray) and the first of the α13-helix. On the bottom part of the figure the resolved sub-ensembles are shown. For $M_{1,\alpha13\ attached}$, cyan spheres represent the selected structures of the main ensemble $M_{1,\alpha13\ near}$ (magenta spheres), for $M_{1,\alpha13\ far}$, the sub-ensemble is shown in blue spheres and for $M_{2,\alpha13\ far}$ in green spheres. Spheres represent the last amino acid of the α-helix. Orange spheres

represent the whole ensemble (around 3000 structures). Structures were aligned to the α12-helix. beginning of the α13-helix.

Monitoring the behavior of the α12 and α13 helices using single-molecule (sm) FRET experiments with Multiparameter Fluorescence Detection (MFD) showed complex dynamics. hGBP1 was labeled in two different schemes, having one dye at the α12-helix (E521C and E533C mutations) and the other at the α13-helix (K582C, M583C). This results in two different samples (E521C/K582C, E533C/M583C) both tracking distance fluctuations and movements of α12/α13 in great detail, since the movement can be fully resolved by FRET derived distances and states (expected distances see **Supplementary Table 11**). Doing so, we resolved a more complex dynamical behavior (see **Figure 42a**) compared to the two-state dynamical equilibrium that has been resolved in the previous study [18]. In these two mutants monitoring the α12/α13-helix movements we resolved a very low FRET distance (see

Supplementary Table 10) on top of the expected distances for the resolved ensembles $M_{x,\alpha13\ near}$. This was done using a lifetime fit of the sm-MFD bursts based on a global fit of the donor only population with a Gaussian distribution fit of the FRET population described in previous literature [21]. Using a fixed width of the Gaussian distribution (σ = 6 Å) led to the limiting state of each species respectively. These distances served as input parameters for the dynamic model (overview of connected methods see **Supplementary Figure 26**) used in the 2d MFD histogram (see **Figure 42a** for E533C/M583C and **Supplementary Figure 22** for E521C/K582C) and in the Photon Distribution Analysis (PDA) (results under **Table 7** and **Supplementary Figure 23**). We interpret the resolved distance as a result of the α13-helix diffusing around, i.e. changing the angle between α12 and α13, where its low FRET distance is the maximum spacing between the labels on α12 and 13-helices (widest angle), defined as $M_{x,\alpha13\ far}$, while the others refer to the position of the α13-helix in the M_1 and M_2 ensembles being near to the α12-helix, defined as $M_{x,\alpha13\ near}$. Testing this by creating a 3d structure with a wide angle between the α12/13-helices got reasonable distances (see **Supplementary Table 11**). Based on a theoretical approach to describe a multiple state system using FRET-lines described previously [22], we applied this principal to our data set (see **Figure 42a**). To sum up, we resolved a structural M_1 ensemble, which contains the states $M_{1,\alpha13\ attached}$, $M_{1,\alpha13\ near}$, $M_{1,\alpha13\ far}$ relating to the position of the α13-helix in respect to the α12-helix, and a structural M_2 ensemble containing $M_{2,\alpha13\ near}$ and $M_{2,\alpha13\ far}$.

The resulting five state system for hGPB1 involved one static state $M_{1,\alpha13\ attached}$ and four states in a dynamic equilibrium, which are both $M_{1,\alpha13\ near}$ and $M_{2,\alpha13\ near}$ states and additionally for each of it a state where the α-13 helix is tilted away to a wide-angle conformation $M_{x,\alpha13\ far}$, as described before. Since $M_{1,\alpha13\ far}$ and $M_{2,\alpha13\ far}$ have approximately the same FRET efficiency, these states were combined to one in the FRET lines approach. This leads to the kinetic model described in **Figure 41**.

$$\overbrace{M_{1,\alpha13\ attached}}^{Static\ fraction} + \overbrace{\underbrace{\left(M_{1,\alpha13\ far} \leftrightharpoons M_{1,\alpha13\ near}\right)}_{Fast\ local\ dynamic\ of\ \alpha13\ of\ M_1} \leftrightharpoons \underbrace{\left(M_{2,\alpha13\ near} \leftrightharpoons M_{2,\alpha13\ far}\right)}_{Fast\ local\ dynamic\ of\ \alpha13\ of\ M_2}}^{Slow\ global\ dynamic\ of\ \alpha12\ switching\ M_1 \leftrightharpoons M_2}$$

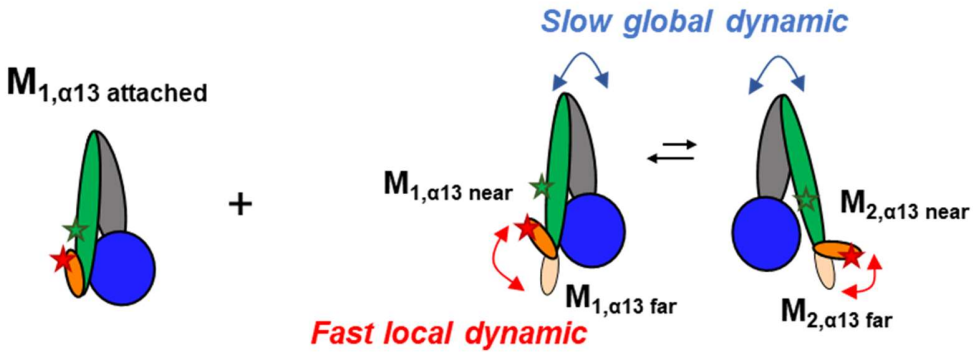


Figure 41: Used linear model resolving the dynamical behavior of the α12/13-helix. The equation shows the linear model used for the PDA. The blue marked region represents the global dynamics of the α13-helix rolling around the MD of the hGBP1. The red marked region represents the fast-local dynamics for the α12-helix. Bottom part of the figure represents the linear kinetic model in a sketch.

A linear model is used since the rates (see **Table 7**) of the diffusing α-13 helix (see **Figure 42**a bottom, red arrows) are short in comparison to the global motion of the α-12 helix (see **Figure 42**a bottom, blue arrows). We will define the diffusion of the α-13 helix as a fast-local dynamic of the M_1 and M_2 states with the α-13 helix switching from a position being near ($M_{x\ near}$) and far ($M_{x\ far}$) from the α-12 helix. The α-12 helix switching its side is a slow global dynamic responsible for the $M_{1,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ near}$ equilibrium. Similar relaxation times were found using FCS [23], correlating High FRET (HF) and Low FRET (LF) bursts (see **Supplementary Figure 24**). Therefore, direct transitions of, for example, $M_{1,\alpha13\ near}$ to $M_{2,\alpha13\ far}$ can be neglected in the linear model used in the PDA. The PDA yields to the majority of the molecules being in $M_{1,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ near}$ exchange, and in $M_{1,\alpha13\ near} \rightleftharpoons M_{1,\alpha13\ far}$ whereas only smaller fractions show the static state of $M_{1,\alpha13\ attached}$ and $M_{2,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ far}$ (see **Table 7**). We used this result as starting parameters for a sub-ensemble TCSPC analysis with a free Gaussian distribution width σ as a fit parameter, and got matching results (see **Table 8** and **Supplementary Figure 25**), which confirms the dynamical characteristics of these mutants. We summarize, that these findings for the α12/α13 behavior of hGBP1 show that the α13 is tumbling back and forth from the α12 defined as a “local” dynamic, whereas the α12 is doing a slower movement, leading to a global change of the molecule. Flipping the side of the α12-helix enables the molecule to enter its M_2 state, which was previously suspected to be important for the protein to build a bridged dimer [24, 25]. In the previous study, the movement of the α13-helix did not lead to a significant change of the FRET efficiency value because of the used label schemes (see **Supplementary Table 11**) and was therefore not identifiable. As a next step in this study, it was possible to determine the importance of these states for the dimerization and oligomerization by looking at farnesylated hGBP1.

Farnesylation locks monomeric hGBP1 in a single state

In order to investigate the impact of farnesylation, the already mentioned farnesylated samples were measured using the same techniques described in the previous section (sm-MFD measurements results under **Supplementary Figure 27**). In direct comparison to non-farnesylated hGBP1 samples from our studies and the ones from the previous study [18], all mutants show a reduction of the local and global dynamical behavior. Exception being the N18C/Q344C$_{farn}$, which displays no dynamics in both the non-farnesylated and farnesylated state, due to the same distance of the dyes in the M_1 and M_2 state. Samples that were labeled in the MD and α12-helix of hGBP1 (Q344C/A496C$_{farn}$ and Q344C/Q525C$_{farn}$) showed a high reduction of the dynamic fraction towards a more static fraction in the M_1 state, resulting in a FRET population nearly on the static line. We assume that here small movements and fluctuations between the MD and α12-helix lead to a slight shift off the static FRET line. This was verified by a lifetime fit of the sm-bursts using a Gaussian distribution fit for the FRET population with a width σ = 10 Å slightly higher than the pure static samples for which one expects a width of σ = 6 Å (see **Supplementary Table 13**). This indicates a higher flexibility in this region of the protein.

In contrast to that, farnesylated hGBP1 variants showing the relation between LGD and GED (N18C/Q577C$_{farn}$) and between α12 and α13 (E533C/M583C$_{farn}$) became completely static (see **Figure 42b**). The $M_{2,\alpha13\ near}$ state and the previously resolved α13 diffusion cannot be observed anymore, meaning that the farnesyl moiety "locks" the hGBP1 molecule in one single state. In case of E533C/M583C$_{farn}$, the molecule is in a locked M_1 ($M_{1,locked}$) state and loses the complex multi-state kinetics that were observed for the non-farnesylated protein (see **Figure 42b**). It is also noteworthy that the measured interdye distance R_{DA} for state $M_{1,\alpha13\ near}$ changed from 36 Å to a shorter distance of 33 Å for E533C/M583C$_{farn}$ (see **Supplementary Table 13**). This indicates that the structure displayed by farnesylated hGBP1 has even tighter binding between the α12/ α13 than the non-farnesylated hGBP1 $M_{1,\alpha13\ near}$ state, being similar to the $M_{1,\alpha13\ attached}$ state. Since the FRET efficiency for this distance is very high (E = 0.94), the acceptor shows photophysical effects like bleaching resulting in a slight shift off the static line. However, we can summarize that the structural exchange of non-farnesylated hGBP1 is not present in the farnesylated monomeric state. A comparison of the measured distances using farnesylated hGBP1 indicates that the $M_{1,locked}$ state is more similar to the crystal structure (pdb: 6k1z, see **Supplementary Table 14**), than the representative $M_{1,\alpha13\ near}$ structure. Summarized, the farnesylation of hGBP1 leads to the monomeric protein being locked in a static state which displays no dynamics at all. This is in contrast to the suspected importance of the dynamics and $M_{2,\alpha13\ near}$ state regarding its role in dimerization, which is why these farnesylated samples were in turn measured while dimerizing.

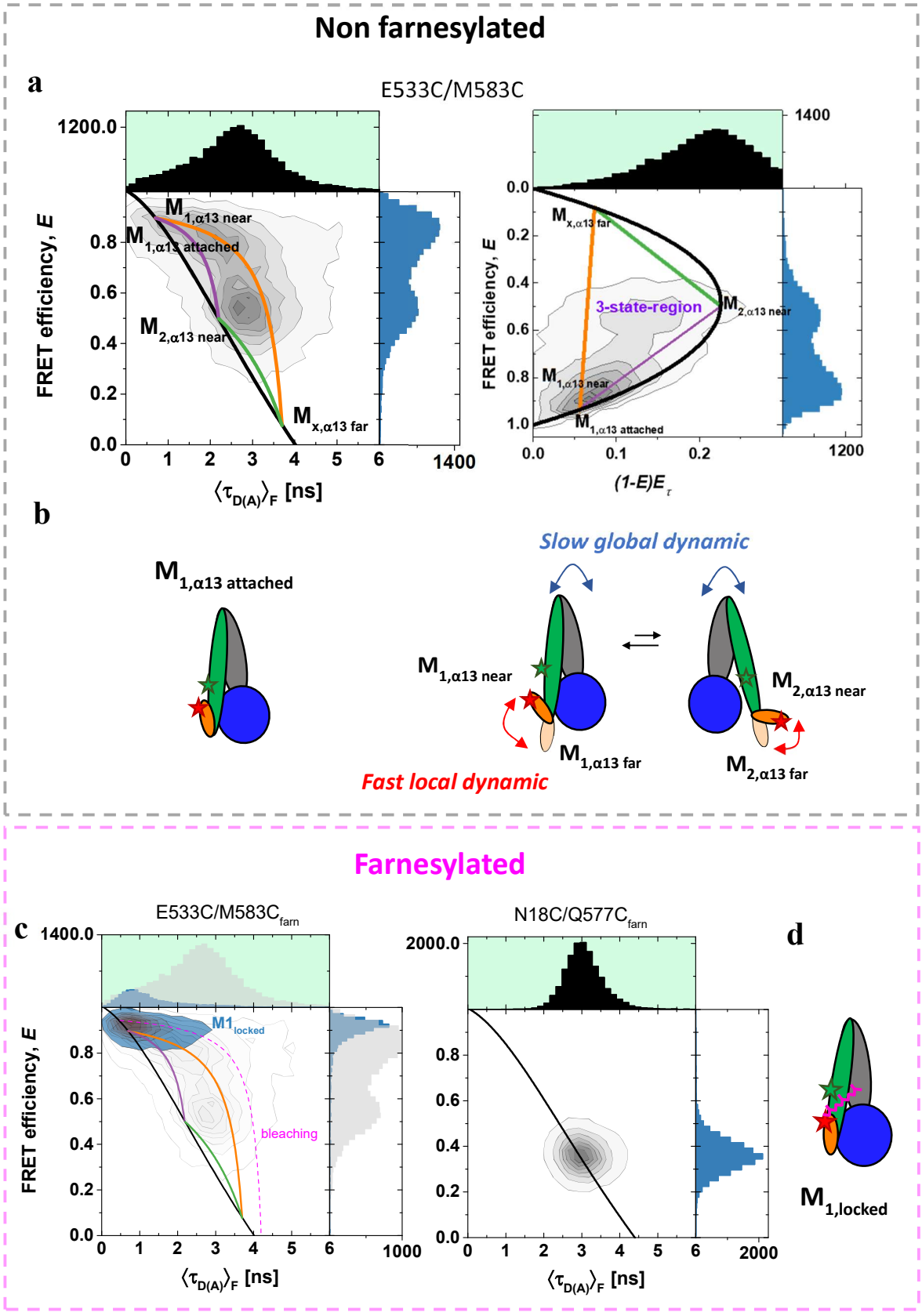

Non farnesylated
a
E533C/M583C
FRET efficiency, E
⟨τD(A)⟩F [ns]
M1,α13 near
M1,α13 attached
M2,α13 near
Mx,α13 far
3-state-region
(1-E)Eτ
b
Slow global dynamic
M1,α13 attached
M1,α13 near
M2,α13 near
M1,α13 far
M2,α13 far
Fast local dynamic
Farnesylated
c
E533C/M583Cfarn
N18C/Q577Cfarn
M1locked
bleaching
d
M1,locked

Figure 42: sm-MFD of hGBP1-Cys9 monitoring the α12 and α13. a (left) 2D sm-MFD histogram of the measurement using the E533C/M583C-sample with the distribution of lifetimes of the donor in presence of an acceptor $\langle \tau_{D(A)} \rangle_f$ (top, black histogram) and the distribution of FRET efficiency values *E* (right, blue distribution). (right) 2D sm-MFD histogram using the FRET efficiency and the difference between the normalized first and second moments of the lifetime distribution as further described in [22]. Bursts were filtered for only double labeled molecules using stoichiometry-, ALEX 2CDE- and TGX-filter. Static and dynamic FRET lines were calculated as described previously [26]. The static FRET line follows 1-(((0.0060*<$\tau_{D(A)}$>$_F^4$)+(-0.0880*<$\tau_{D(A)}$>$_F^3$)+(0.4122*<$\tau_{D(A)}$>$_F^2$)+0.3772*<$\tau_{D(A)}$>$_F$-0.0220)/4.0000), the dynamic line for $M_{1,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ near}$ 1-(0.7000*2.2000/(4.0000*(0.7000+2.2000-((0.0000**<$\tau_{D(A)}$>$_F^3$)+(-0.0000**<$\tau_{D(A)}$>$_F^2$)+1.9608**<$\tau_{D(A)}$>$_F$+-2.1818)))), for $M_{1,\alpha13\ near} \rightleftharpoons M_{x,\alpha13\ far}$ 1-(0.7000*3.7000/(4.0000*(0.7000+3.7000-((0.0000**<$\tau_{D(A)}$>$_F^3$)+(-0.0000**<$\tau_{D(A)}$>$_F^2$)+1.8455**<$\tau_{D(A)}$>$_F$+-3.1306)))) and for $M_{2,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ far}$ 1-(2.2000*3.7000/(4.0000*(2.2000+3.7000-((0.0000**<$\tau_{D(A)}$>$_F^3$)+(-0.0000**<$\tau_{D(A)}$>$_F^2$)+1.2362**<$\tau_{D(A)}$>$_F$+-0.8771)))). FRET lines are indicated as black for the static FRET line and colored for dynamic FRET lines, with each color representing different transitions (purple: $M_{1,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ near}$, orange: $M_{1,\alpha13\ near} \rightleftharpoons M_{1,\alpha13\ far}$, green: $M_{2,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ far}$). Magenta marked region indicates the 3-state dynamic population. All results can be viewed under **Table 8** and **Supplementary Table 10**.

Table 7: PDA of E533C/M583C and E521C/K582C. Resulting fractions for the E533C/M583C and E521C/K582C samples using a linear model with one static and 3 dynamic states. Dynamic fractions and resulting relaxation times can be found under **Supplementary Table 12**.

	Fraction [%]				**Rate [1/s]** rate state 1 to state 2 rate state 2 to state 1		
Species	$M_{1,\alpha13\ attached}$	$M_{1,\alpha13\ near}$	$M_{2,\alpha13\ near}$	$M_{x,\alpha13\ far}$	$M_{1,\alpha13\ near} \leftrightharpoons M_{2,\alpha13\ near}$	$M_{1,\alpha13\ far} \leftrightharpoons M_{1,\alpha13\ near}$	$M_{2,\alpha13\ near} \leftrightharpoons M_{2,\alpha13\ far}$
E533C/M583C	22	44	18	16	0.8 1.5	25 25	20 20
E521C/K582C	18	38	19	25	0.7 1.2	13 16	10 10

Table 8. Sub-ensemble TCSPC fitting of E533C/M583C and E521C/K582C. σ indicates the width of the Gaussian distribution, which is greater than 6 Å for dynamic equilibria.

	Dynamic analysis (*σ* fitted)							
Species	$M_{1,at}$		$M_{1,\alpha13\ near} \leftrightharpoons M_{2,\alpha13\ near}$		$M_{1,\alpha13\ far} \leftrightharpoons M_{1,\alpha13\ near}$		$M_{2,\alpha13\ near} \leftrightharpoons M_{2,\alpha13\ far}$	
	R_{DA} [Å]	σ [Å]	R_{DA} [Å]	σ [Å]	R_{DA} [Å]	σ [Å]	R_{DA} [Å]	σ [Å]
E533C/M583C	35 (20%)	6	43 (48%)	17	51 (30%)	49	64 (2%)	12
E521C/K582C	29 (10%)	6	42 (40%)	15	58 (40%)	47	69 (10%)	13

Both dimerization and hydrolysis are required to unlock farnesylated hGBP1

As described previously, the dynamic equilibrium of the M_1 and M_2 states could not be observed in the monomeric state of farnesylated hGBP1. However, since we suspect that the dynamics play a major role during the activity of the protein, the samples were measured while adding different nucleotides and varying the protein concentration using wild type farnesylated, non-labeled hGBP1 in order to shift the equilibrium towards dimerization (overview see **Table 9**). As test systems, we have chosen E533C/M583C$_{farn}$ and N18C/Q577C$_{farn}$ (see **Figure 43** and **Supplementary Figure 28**), since in these samples the most dramatic changes are expected after dimerization. For a comparison of the dimerization of the non-farnesylated hGBP1 we have performed the same measurements using several mutants (see **Figure 43** and **Supplementary Figure 29**). In **Figure 43** we specifically tested addition of the nucleotide GDP-AlF$_x$ with different hGBP1 concentrations (classification of nucleotide and overview of used nucleotides see **Table 9**). GDP-AlF$_x$ has been shown before to allow the protein to open up (meaning to detach α12/α13 from the GED and MD and eventually to stretch out to like 180°), bind to membranes and polymerize [27, 28]. For the non-farnesylated hGBP1 it could be seen that the opening and dimerization occurred on a time scale not accessible for confocal sm-MFD (**Figure 43a**) which is around one minute. Hence, we interpret that the equilibrium is highly shifted towards the dimer state of the hGBP1, which was confirmed by other mutants, where only a very minor fraction is still in a monomeric state at concentrations of 0.1 µM (see **Supplementary Figure 28**). We can conclude that for the non-farnesylated hGBP1 we have an intrinsic flexibility already in the monomeric state, as seen before already, which after dimerization is dramatically reduced.

For the farnesylated hGBP1, we first added GDP-AlF$_x$ under single molecule conditions (**Figure 43b**) and observed no change to the measurement of farnesylated hGBP1. As these measurements were done with protein concentrations of approximately 20 pM, we can safely assume no dimerization can take place. Hence, the anchor is still fixed and the molecule is in its $M_{1,locked}$ state, showing no dynamic behavior at all. In the next step, 0.5 µM of farnesylated wild type hGBP1 was added. As farnesylated hGBP1 polymerizes irreversibly in the presence of GDP-AlF$_x$, the added amount of farnesylated wild type hGBP1 was kept low to avoid polymerization as much as possible, while still allowing for dimerization. The sm-MFD measurement revealed that under these conditions the monomeric dynamics as observed for non-farnesylated hGBP1 (see section: **Monomeric non-farnesylated hGBP1 exhibits extensive dynamics**) are active again. The 2D *lifetime-efficiency* histogram shows a complex dynamic with $M_{1,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ near}$, $M_{1,\alpha13\ near} \rightleftharpoons M_{1,\alpha13\ far}$ and $M_{2,\alpha13\ near} \rightleftharpoons M_{2,\alpha13\ far}$. The same model as before was used for a PDA with an additional static fraction at a very low FRET efficiency ($M_{x,\alpha13\ far}$) to satisfy the already stable dimer state (see **Supplementary Figure 30**, **Supplementary Table 16**), resulting in the same distances and dynamics. As a result, the fraction of molecules showing a dynamical monomeric behavior are lower and a high static fraction of $M_{x,\alpha13\ far}$ appeared. In a similar way, the N18C/Q577C$_{farn}$ sample was tested after addition of GDP-AlF$_x$ (see **Figure 43c** and **Supplementary Figure 28**) and GTP (see **Supplementary Figure 31Error! Reference source not found.**), respectively. It can be seen, that under sole addition of GDP-AlF$_x$ a diffusion time of the molecule of τ_d = 1.32 ms was obtained, whereas further adding wild type resulted in a diffusion time at the beginning of the measurement of τ_d = 1.82 ms, which results in a factor of 2.5 for the hydrodynamic volume (see **Supplementary Figure 32**). For that, we conclude that dimerization occurred. At later stages of the measurement, oligomers with big and heterogenic diffusion times were found as rare events for E533C/M583C$_{farn}$, but as multiple events for N18C/Q577C$_{farn}$ (see **Supplementary Figure 33**). In the case of N18C/Q577C$_{farn}$ only three states were expected, which is M_1, M_2 and a no-FRET species, where large distances between the FRET lables due to opening

of the GED (α12/13) is expected. The species fractions of these three states were fitted in a sub-ensemble TCSPC fit in dependency of time (see **Supplementary Table 17** and **Supplementary Figure 34**). It can be seen that the FRET active species decays exponentially with a lifetime of τ = 460 ± 100 s. We can deduce that this formation of a no-FRET species is the expected opening of α12/13 described above.

We conclude that only adding GDP-AlF_x, which is known to be required for protein activity, does not change the behavior of the molecule at all. With additional wild type hGBP1 in the µM-range we see a return of the dynamics previously only observed for non-farnesylated hGBP1 coupled with an increase of diffusion time. Therefore, we deduce, that both nucleotide binding and dimerization are required for farnesylated hGBP1 to leave its "locked" M_1 state. With M_1 and M_2 once again being displayed by the protein, it is able to form dimer interfaces with the GED, potentially leading to the bridged dimers. These findings also answer the question raised in the previous hGBP1 study [18], as this shows, that the flexibility is induced via dimerization for farnesylated hGBP1. After a few minutes (see **Supplementary Figure 34**) the observed occurrences of FRET and no-FRET populations stabilize, and dynamic events decrease dramatically. Based on these observations, we deduce that the final state of the irreversible activity with GDP-AlF_x is reached and that this includes the possible formation of oligomers. However, based on the observed diffusion times and occurence of bigger structures, it is likely that the most abundant species is still the hGBP1 dimer. We would therefore observe a bridged dimer, which is able to adapt a open conformation, similar to what is expected within the oligomer. We could also see that contrary to the non-farnesylated hGBP1 the formation of the dimer is slower.

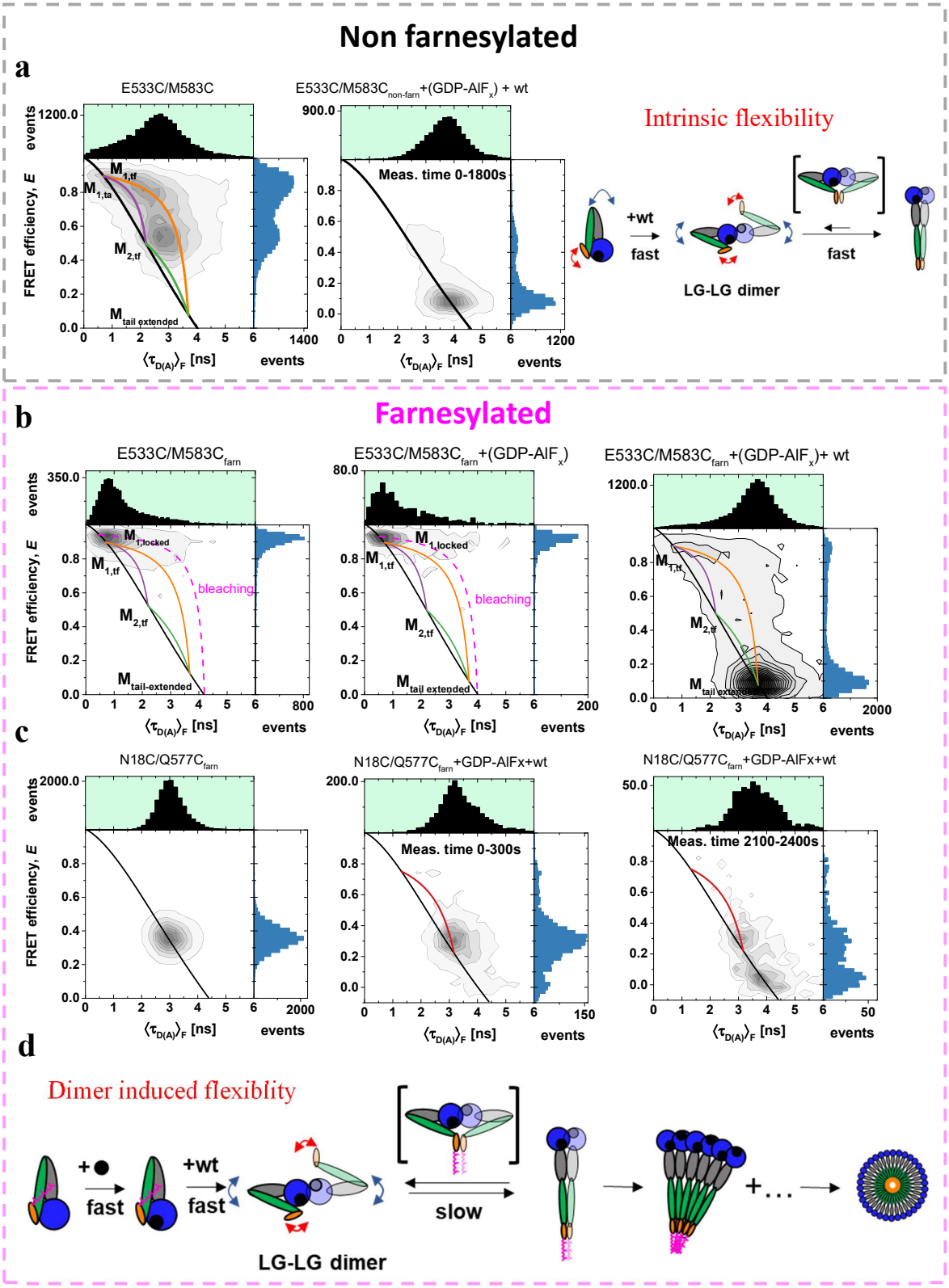


Figure 43: Nucleotide studies using GDP-AlF$_x$. a 2D *lifetime-efficiency* histogram of E533C/M583C of the non-farnesylated hGBP1 without (left) and under the addition of GDP-AlF$_x$ and unlabeled wild type (0.1 µM) (middle). The FRET efficiency values immediately drop to very

low values, with only a small fraction of molecules not in the dimeric state. Derived model is shown as a sketch on the right-hand side. **b** 2D *lifetime-efficiency* histogram of E533C/M583C of the farnesylated hGBP1 without (left), under the addition of GDP-AlF_x (middle) and under the addition of GDP-AlF_x and unlabeled wild type. **c** 2D Donor-*Lifetime* -FRET *Efficiency* histogram of N18C/Q577C of the farnesylated hGBP1 without (left), after the addition of GDP-AlF_x and unlabeled wild type at the beginning of the measurement (middle) and at the end of the measurement (right). **d** Derived model shown as a sketch.

Table 9: Overview of the used nucleotides, their biological relevance and their observed effect on the measurement with farnesylated hGBP1 samples.

Nucleotide	Measuring conditions	Biological relevancy	Effect on measurement
GTP	2 mM GTP 0.1, 5 µM wild type	Natural substrate	After time delay, fluorescence disappears for some time, then returns
GMP	250 µM GMP 10 µM wild type	Natural product, binding strength is similar to GTP	No change
GTPyS	250 µM GTPyS 10 µM wild type	Simulates GTP bindung, not hydrolysable	No or little change
GDP-AlF_x	200 µM GDP-AlFx 0.1, 0,5, 1, and 10 µM wild type	Simulates state between GTP binding and hydrolysis (GTP-GDP)	Concentration dependent change of detected fluorescence

Polymerization results in a static and elongated state of hGBP1

As a final step we studied farnesylated hGBP1 under conditions, where oligomers could be formed by adding more (10 µM) wild type farnesylated hGBP1, providing typical conditions for oligomerization [27, 29]. Measuring complexes of this size is a challenging task for diffusional measurement techniques like sm-MFD since the oligomers are gravitating to the surface, whereas the focus is typically a few µm above. This limits the measurement time to a point where the majority of oligomers is on the surface, which is typically around 30 minutes. Therefore, all five farnesylated samples (E533C/M583C$_{farn}$, N18C/Q577C$_{farn}$, N18C/Q344C$_{farn}$, Q344C/A496C$_{farn}$ and Q344C/Q525C$_{farn}$) were additionally measured using imaging microscopy, to double check and verify the results (see **Supplementary Figure 35**). To filter out the oligomers from monomers/dimers, only bursts were used that had a count rate much higher than the average one. Additionally, FCS showed correlation curves indicating that aggregates with high diffusion times are contained in the sample (**Supplementary Figure 36**). Moreover, the distribution of burst durations also indicated that bursts are contained which have a higher burst duration than a monomer would have (**Supplementary Figure 36**). All variants (except N18C/Q344C$_{farn}$) showed a large shift towards higher inter dye distances R_{DA} (see **Supplementary Table 15**) leading to very low FRET efficiencies. Comparing these results of confocal measurements to the imaging experiments led to the same results (see **Supplementary Table 15** and **Supplementary Figure 37**). Since an opening of farnesylated hGBP1 upon oligomerization is expected [18, 25, 27-29] the measurements confirm these expectations very well. In order to further quantify this, a simple model was assumed where the crystal structure 6k1z from PDB data base was taken and the α12- and α13 helix were manually changed to get a fully opened molecule. This model was used to compare the distances that have been measured. This approximation of how a fully opened hGBP1 could look like, fit the observed distances very well (see **Figure 44**). In recent experiments, the variant N18C/Q577C$_{farn}$ was measured to determine the distance between N- and C-terminus within the oligomer, amounting to around 25-28 nm [25, 29]. This range is far away from the measurable FRET distances using the dye combination Alexa488/Alexa647 with a Förster Radius of 52 Å. To still be able to compare this expected distance range with FRET derived distances, the projected distance of each measured mutant to the 18-583 distance was calculated using the 3D model of the fully opened hGBP1 derived before, including a complete widening of the angle between α12 and α13. Following basic geometry and taking the FRET derived distance as the absolute value of a vector $\left|\overrightarrow{R_{DA}}\right|$, where $\overrightarrow{R_{DA}}$ is the vector connecting the mean positions of the corresponding dyes using a 3D structure, one can compare measured distances to the fully elongated model. For a pair of vectors *i* and *j* starting at the same point, the projection of vector $\overrightarrow{R_{DA,i}}$ to vector $\overrightarrow{R_{DA,J}}$ is defined as:

$$\overrightarrow{R_{DA,i}}_{,DA,j} = \frac{\overrightarrow{R_{DA,i}}\,\overrightarrow{R_{DA,J}}}{\left|\overrightarrow{R_{DA,J}}\right|^2}\,\overrightarrow{R_{DA,J}} \ .$$

Using this for each measured distance $\left|\overrightarrow{R_{DA}}\right|$ the projection to $\overrightarrow{R_{18-583}}$ was calculated (see **Supplementary Table 15**). Adding the projected distances for all measured distances leads to the distance $\left|\overrightarrow{R_{18-583}}\right|$. Addition of the FRET derived projected distances resulted in $\left|\overrightarrow{R_{18-583}}\right|_{\text{FRET}}$ = 260±30 Å, compared to the distance for the mean dye positions suggested from the 3D model of the fully opened hGBP1 $\left|\overrightarrow{R_{18-583}}\right|_{\text{model}}$ = 251 Å. This agreement confirms the assumed opening of hGBP1 upon oligomerization and further proves complete opening both between MD and α12 and between α12/α13.

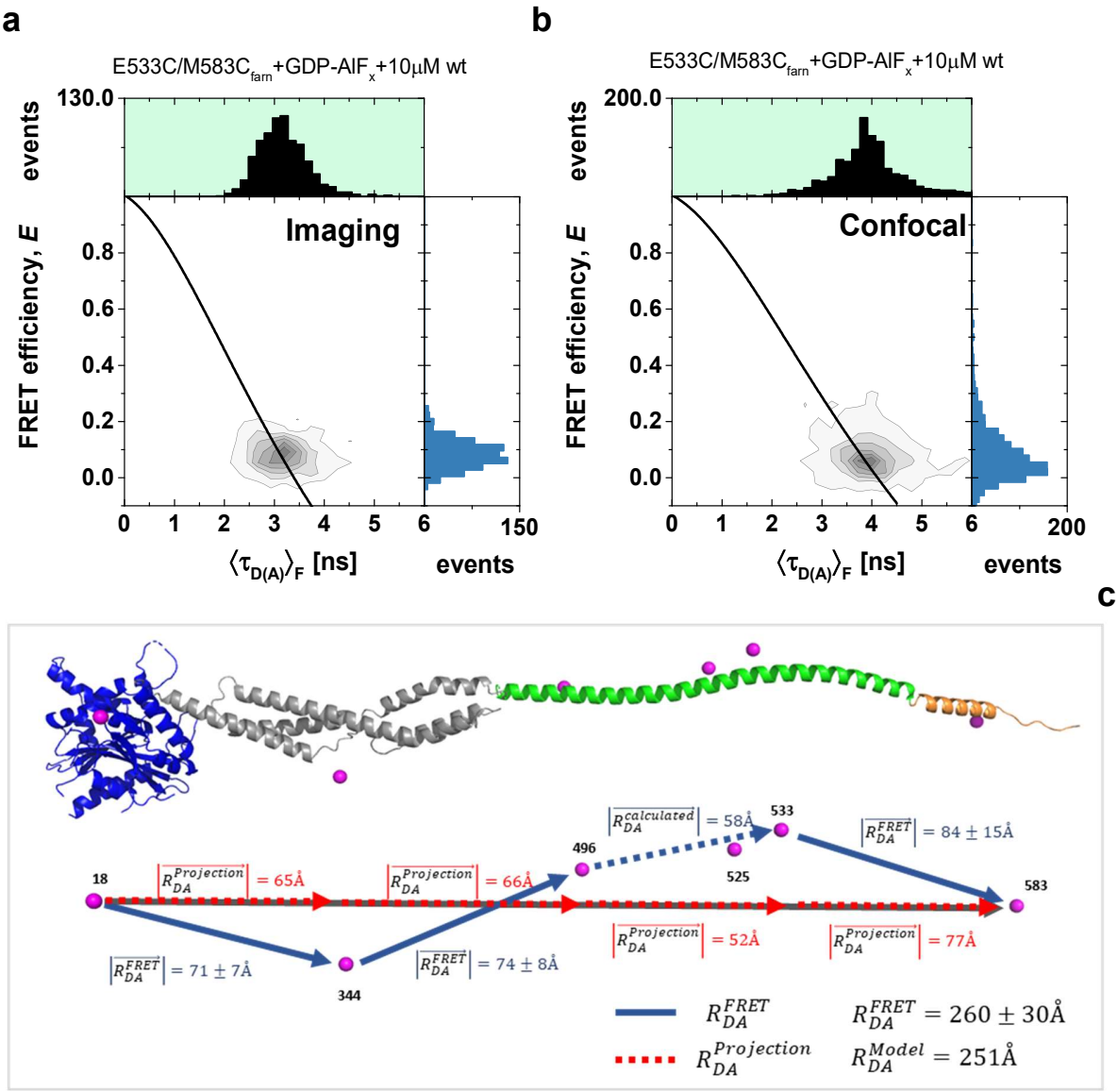


Figure 44: Resolving the oligomeric state of farnesylated hGBP1. **a** 2d-MFD imaging measurement of the E533C/M583C$_{farn}$ sample and **b** the corresponding confocal sm-MFD measurement. In both cases, a nucleotide (GDP-AlF$_x$) and a high concentration of farnesylated, unlabeled wild type (10µM) was added, with used concentrations of 300 µM $AlCl_3$, 10 mM NaF and 200 µM GDP. Black line represents the static FRET line. Bursts were filtered for only double labeled molecules using stoichiometry-, ALEX 2CDE- and TGX-filter, in case of imaging sufficient donor photons after donor excitation and sufficient acceptor photons after acceptor excitation. **c** (top) 3d-model of hGBP1 with opened α12-helix (green) and α13-helix at widest angle (orange) manually modified based on pdb 6k1z. (bottom) Gray arrow below red dashed arrows represents the vector from the mean position of the dye at amino acid N18 to M583. On top of this, red dashed arrows indicate the projection of the blue vectors. The absolute value of the blue arrow is R_{DA}, measured by FRET. By estimating the angle between red dashed and blue vector, the absolute value of red dashed vector can be calculated. The sum of red dashed vectors corresponds to the FRET derived distance of the mean positions of the dye from N18 to M583 (260±30 Å), which can be compared to the value the shown 3d structure suggests (251 Å).

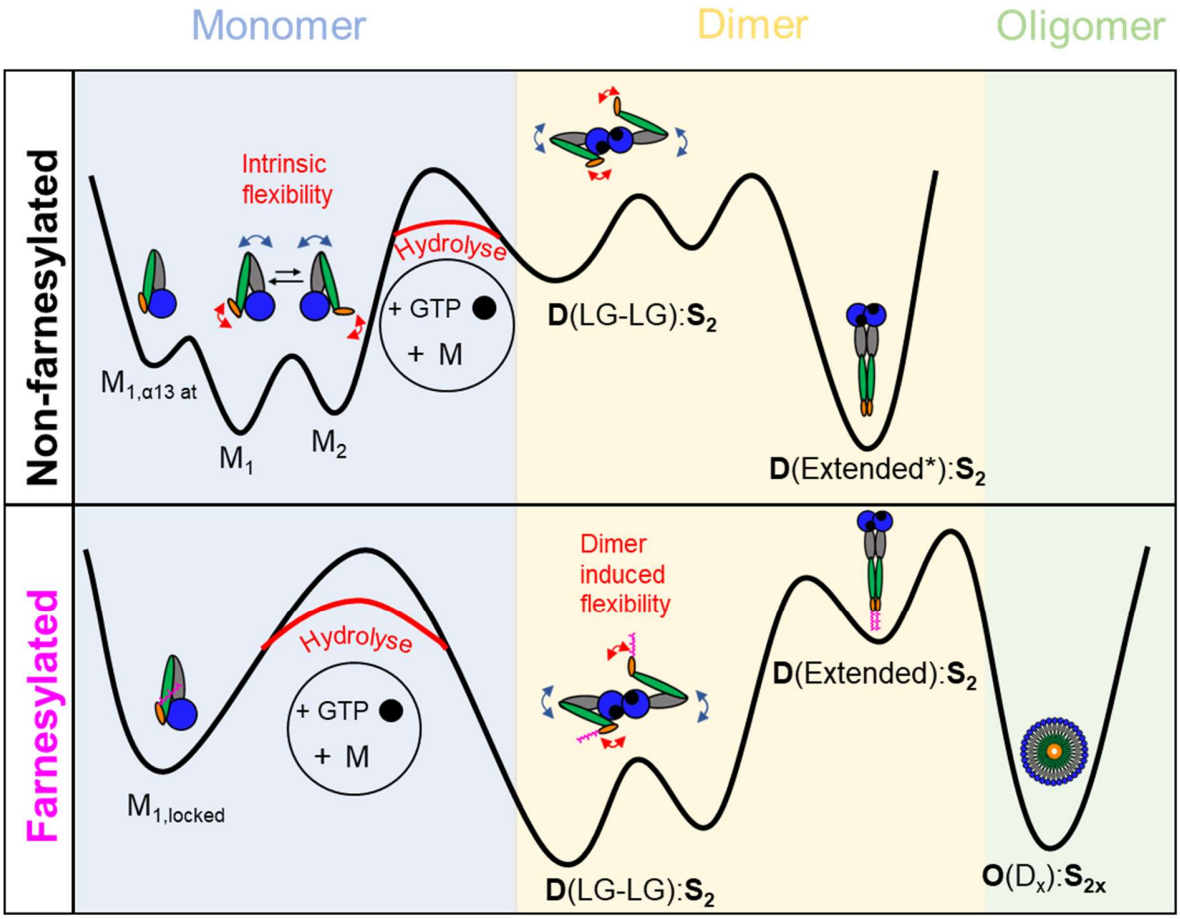


Figure 45: Potential global pathway of the non- and farnesylated hGBP1 in comparison. Blue background indicates the monomeric, orange the dimeric and green the oligomeric region. (Top) Pathway of the farnesylated hGBP1. For farnesylated hGBP1 no dynamic exchange for the monomeric state was found, even after adding a nucleotide. Increasing the concentration enables the possibility to form dimers, the same dynamical pathway of the non farnesylated hGBP1 was found in the LG-LG dimer. After a bridged dimer is formed with additional GED-GED contact, the molecule can open completely (stretch out to 180°) and potentially oligomerize in case of higher concentrations. (Bottom) Pathway of non-farnesylated hGBP1 with a static state $M_{1,at}$ and dynamic interactions of M_1 and M_2 (blue arrows) implying the additional diffusion of the α13-helix (red arrows).

Discussion

Using specifically labeled samples for sm-FRET studies, it was possible to unravel many structural and dynamical dependencies of hGBP1 based on singular farnesylation, crucial for its activity cycle. This was possible due to the connection of different analysis methods on different stages of the protein (see **Supplementary Figure 26**). As mentioned before, hGBP1 was shown to display extensive conformational changes during its activity, mainly dimerization, membrane binding and polymerization [27, 28, 30]. Based on previous research using non-farnesylated hGBP1, it was suspected that two different monomeric states (M_1, M_2) are responsible to form dimer interfaces [18]. We could improve on this work by adding more data, proving the existence of these two states and refining the expected structure with an additional characterization of the dynamic of the α12/13-helices of the hGBP1.

For farnesylated hGBP1, which is supposed to be the biologically active one capable of membrane binding and polymerization, literature has postulated a much tighter structure than for non-farnesylated hGBP1 [13, 14]. In this work we could show with quantitative data, that farnesylated hGBP1 samples displays no dynamics previously found in non-farnesylated samples and its interaction between the α13 and α12 helices. As assumed in previous research [14], the farnesyl moiety locks the protein and impedes dynamics, acting as a controlling factor to what dynamics the protein displays in monomeric form. Dimerization was studied by adding a nucleotide analogue (GDP-AlF_x). Curiously, under single molecule conditions, this nucleotide addition yielded no effect on the behavior of the protein. Effects could only be observed when adding enough protein to allow for dimerization. Under these conditions, the previously determined dynamics could also be observed for farnesylated hGBP1 samples. Taking these observations into account, farnesylated hGBP1 requires both an appropriate nucleotide and the dimerization to detach the farnesyl moiety and leave its locked state (see **Figure 45**, bottom region). As the first step of dimerization for hGBP1 is the formation of a LG-LG interface [30], the GEDs would be on opposite ends of the protein. We conclude that the global dynamic between a $M_{1,\alpha13\ near}$ and $M_{2,\alpha13\ near}$ state is required for the GEDs to face each other, while the fast local dynamic of the α13 acts as a searching tool to form a GED-GED interface with its dimer counterpart to form a bridged dimer. Therefore, dimer induced flexibility [18] combined with the necessity of complex dynamics are required to form crucial dimer interfaces for hGBP1. Additionally, it could be shown that the farnesylation significantly slows the dimerization process, as no kinetics could be seen with equivalent experiments with non-farnesylated hGBP1 samples. Contrary to this, the non-farnesylated hGBP1 possesses these dynamics already in the monomeric state (see **Figure 45**, top region), hence it has an intrinsic flexibility. It could be seen that dimerization occurs fast and that the equilibrium is shifted highly towards the open dimer.

The final state of these dimerization experiments for both non-farnesylated and farnesylated hGBP1 with GDP-AlF_x displayed the same distances. Labels monitoring the α13-helix of the hGBP1 indicate a high increase in the distance in its dimeric (and oligomeric for the farnesylated hGBP1) state. For farnesylated hGBP1, it is expected to form higher ordered polymers with an extended structure based on localization and cryogenic electron microscopy experiments [25, 29]. By assuming an extended structure model and simulating mean dye positions on it, we could compare our quantitative distances to find good agreement with a completely opened structure of hGBP1 while in the oligomeric state. This involves both the complete detachment of α12 from the MD and α13 from the α12 leading to an opening of the molecule. As the non-farnesylated hGBP1 displays exactly the same distances expected of an opened structure as well, it is most probable that it also opens completely during dimerization. Due to the nature of the data it cannot be excluded, that non-farnesylated hGBP1 assumes an angular structure instead, possibly explaining the faster kinetics during dimerization. We denote that this feature of the molecule is already

present in the monomeric state of the non-farnesylated hGBP1, indicating that the structure is highly connected to the possible pathway of the molecule.

All in all, we could show the necessity and extensive impact of farnesylation on both the dynamics and structure of hGBP1. In this case, the farnesylation acts as a control to lock dynamics necessary for dimerization to only occur after nucleotide binding and a partner protein being present at concentrations high enough for dimerization. This is important to note as this safety mechanism prevents hGBP1 to disappear at low concentrations from the cytosol by attaching to membranes. Furthermore, it allows the subsequent formation of higher ordered structures and membrane binding by stabilizing a completely opened structure.

Methods

Protein sample preparation

Protein expression, farnesylation and purification were performed as described in previous work [30-32]. For site-specific labeling of hGBP1, first all native cysteine residues were mutated (i.e., C12A, C82A, C225S, C235A, C270A, C311S, C396A, C407S, C589S), as used before [18]. The addition of two cysteines allowed for site specific labeling of those non-farnesylated hGBP1 variants. Reversal of the C-terminal C589S mutation allowed for site-specific farnesylation, resulting in farnesylated hGBP1variants (see **Supplementary Figure 38** and **Supplementary Figure 39**). All hGBP1 variants were labeled with Alexa 488 maleimide, purified by anion exchange chromatography and then labeled with Alexa 647 maleimide. Samples were tested for their enzymatic activity both before and after labeling (see **Supplementary Figure 38**). The total protein concentration in a measured sample varied depending on the desired hGBP1 state that is to be measured. For monomer measurements, the labeled sample was measured in single molecule conditions (approximately 20 pM). For dimerization and oligomerization experiments, additional wild type hGBP1 was added to the measurement to achieve required protein concentrations, which amounted to 0.1-1 µM for dimerization and 10 µM for oligomerization. Dimerization and oligomerization were induced by the addition of 200 µM GDP in the presence of AlF_x in buffer containing 50 mM Tris-HCl, 150 mM NaCl, 5 mM $MgCl_2$, 300 µM $AlCl_3$ and 10 mM NaF at pH 7.9. For oligomerization samples to be measured with confocal microscopy the workflow was analogous to previous work [29]. In short, a typical oligomerization sample was reacted for 15 min and then diluted 1:1000 in buffer containing 50 mM Tris-HCl, 150 mM NaCl, 5 mM $MgCl_2$, 300 µM $AlCl_3$, 10 mM NaF and 200 µM GDP before applying it to a Nunc™Lab-Tek™II Chamer Slide™ (Thermo Fisher Scientific, USA). Large polymer structures would appear on the surface after 10-40 min of incubation. Additional information regarding the enzymatic activity of the used hGBP1 samples is given in the supplementary information.

Single-molecule experiments

Single-molecule Multiparameter Fluorescence Detection (sm-MFD) experiments were performed on a home-built setup based on an Olympus IX70 inverted microscope as described in [33]. Further details can be found in the SI-section “Instrument used for confocal fluorescence spectroscopy”. Double labeled molecules were selected via stoichiometry values using two excitation sources in a PIE scheme as described in literature [17, 34]. Burst selection was done using a time threshold and a minimum number of photons [33]. Correction factors for intensity based confocal MFD were calculated using methods described in [35]. In short, fluorescence signal is corrected for crosstalk/donor leakage, background, different excitation flux, quantum yields and the ratio of detection efficiency. To determine the ratio of detection efficiency each spectrum of the optical components was measured and overlayed to the donor and acceptor emission spectrum, respectively. The ratio of the overlap integral determines the relative detection efficiency.

Fluorescence decay analysis

Ensemble Time-Correlated-Fluorescence-Decays (eTCSPC) were taken using a Fluotime 200 (PicoQuant, Germany). For excitation, a white light laser source was used (SuperK, NKT Photonics) operated at λ=485 nm. Detection was done under magic angle conditions. For Instrument Response Function (IRF) a Ludox solution was used. Sub-ensemble Time-Correlated-Fluorescence-Decays (seTCSPC) were extracted from sm-MFD measurements via histogram of

all photons which are contained in the selected bursts. Donor in absence and presence of an acceptor were fitted globally using software "ChiSurf" described in [21] and available on GitHub (https://github.com/Fluorescence-Tools/chisurf). eTCSPC data was used for structural modeling, seTCSPC was used for dynamical analysis (for overview, see Supplementary Figure 26).

Distance determination and structural modeling

Distance determination was done using eTCSPC. Gaussian distribution model was applied to double labeled sample and fitted globally with the single labeled donor only sample. For each FRET sample, two Gaussian distributions were used, leading to distances for each monomeric state, M_1 and M_2 of the hGBP1, where the major fraction is the one with a higher amplitude. Due to linker dynamics [36] a width of σ=6 Å was used. Errors in the distance were estimated using standard error propagation of different sources:

$$\delta_{R_{DA}} = \sqrt{\delta_{AV}^2 + \delta_{\kappa}^2 + \delta_{Reference}^2 + \delta_{statistical}^2}$$

with error contributions of accessible volume model δ_{AV}^2, the orientation factor of the dyes δ_{κ}^2, the reference of the donor only sample $\delta_{Reference}^2$, and the statistical noise of the data $\delta_{statistical}^2$. Estimating errors was done following [18]. For structural modeling each distance for M_1 and M_2 state of the hGBP1 of the ensemble was calculated using Accessible Contact Volume (ACV) approach [37], defining the volume of the possible position of the dye including a contact area on the surface of the biomolecule. For simulation, standard linker parameters were used which is a linker length=22 Å, linker width= 2.5 Å, and a single radius of 4 Å. Simulated distances $R_{DA,sim}$ were compared to experimentally resolved distances $R_{DA,exp}$ using χ^2-approach:

$$\chi_{FRET}^2 = \sum_i \left(\frac{R_{DA,exp}^{(i)} - R_{DA,sim}^{(i)}}{\delta_{R_{DA}}^{(i)}} \right)^2$$

χ^2 values of additional FRET pairs were added to the previously measured χ^2, and merged with EPR and SAXS to finally discriminate structures as described in section "Monomeric non-farnesylated hGBP1 exhibits extensive dynamics" and previously [18].

Photon distribution analysis

PDA was done using approaches described in [38]. Time Windows (TW) sizes of typically 1, 2 and 3 ms were chosen. For fitting a standard Levenberg-Marquardt algorithm was used. A linear model was used with dynamic and static fractions as described in section "Monomeric non-farnesylated hGBP1 exhibits extensive dynamics".

FRET-lines

FRET lines were calculated as described in [22, 39]. In short, static FRET lines describe the ratio between the FRET efficiency *E* and the lifetime of the donor in presence and absence of the acceptor as follows:

$$E_{static} = 1 - \frac{\langle \tau_{D(A)} \rangle_f}{\tau_{D(O)}}$$

with $\langle \tau_{D(A)} \rangle_f$ denoting the fluorescence weighted lifetime of the donor in presence of an acceptor and $\tau_{D(O)}$ the species weighted lifetime. After correction for linker dynamics, molecules show

FRET efficiency and lifetimes on the static FRET line, whereas molecules with dynamic are off the static FRET lines. In case of dynamics occurring while the molecule is in the focus, the fluorescence weighted lifetime $\langle \tau_{D(A)} \rangle_f$ is biased towards higher lifetimes due to higher photon weighting of longer lifetimes. Therefore, population with dynamics are shifted towards higher lifetimes off the static FRET line on a dynamic FRET line connecting the two limiting states. In case of a multi-state FRET system as measured in section “Monomeric non-farnesylated hGBP1 exhibits extensive dynamics”, FRET populations are in between dynamic FRET lines because of the implication of more than 2 FRET states. This kinetic network was analyzed using multiple methods, as PDA, seTCSPC and fFCS, to resolve the involved limiting states.

Data availability

Following material is available at Zenodo: Experimental data (10.5281/zenodo.7458187): (1) fluorescence decays recorded by eTCSPC used to compute distances for structural modeling (2) sm-MFD: raw data, calibtration measurements and burst selection (3) imaging raw data and (4) sub-ensemble fluorescence decays. Scripts for structural modeling and selected ensemble can be found at (10.5281/zenodo.6565895). Structure models of hGBP1 based on experimental restraints were deposited to PDB-Dev () using the FLR-dictionary extension (developed by PDB and the Seidel group) available on the IHM working group GitHub site (https://github.com/ihmwg/FLR-dictionary).

Acknowledgment

We thank Thomas-Otavio Peulen for groundwork in previous hGBP1-study and helpful input, Mykola Dimura for doing the screening to find the most informative FRET pairs and Noah Salama for assisting sm-MFD measurements of non-farnesylated hGBP1. J.F. acknowledges support of the International Helmholtz Research School of Biophysics and Soft Matter (BioSoft). C.S. acknowledges support from the European Research Council (ERC) through the Advanced Grant 2014 hybridFRET (Grant No. 671208).

Author contributions

JF and PL wrote the initial draft of the manuscript. PL prepared the samples and measurement conditions. JF performed sm-confocal fluorescence experiments and analyzed data. PL performed confocal fluorescence imaging spectroscopy and analyzed data. JF and PL interpreted data. CH and CAMS designed the research and supervised the project. JF, PL, CH and CAMS interpreted the data and wrote the manuscript.

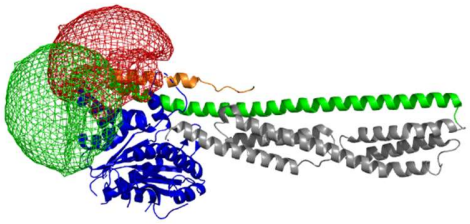
1. Wang, M. and P.J. Casey, *Protein Prenylation: Unique Fats Make Their Mark on Biology.* Nature Reviews Molecular Cell Biology, 2016. **17**(2): p. 110-122.
2. Palsuledesai, C.C. and M.D. Distefano, *Protein Prenylation: Enzymes, Therapeutics, and Biotechnology Applications.* ACS Chemical Biology, 2014. **10**(1): p. 51-62.
3. Berndt, N., A.D. Hamilton, and S.M. Sebti, *Targeting Protein Prenylation for Cancer Therapy.* Nature Reviews Cancer, 2011. **11**(11): p. 775-791.
4. Honkala, A.T., D. Tailor, and S.V. Malhotra, *Guanylate-Binding Protein 1: An Emerging Target in Inflammation and Cancer.* Frontiers in Immunology, 2020. **10**.
5. Lubeseder-Martellato, C., et al., *Guanylate-Binding Protein-1 Expression Is Selectively Induced by Inflammatory Cytokines and Is an Activation Marker of Endothelial Cells During Inflammatory Diseases.* The American Journal of Pathology, 2002. **161**(5): p. 1749-1759.
6. Pilla-Moffett, D., et al., *Interferon-Inducible Gtpases in Host Resistance, Inflammation and Disease.* Journal of Molecular Biology, 2016. **428**(17): p. 3495-3513.
7. Tretina, K., et al., *Interferon-Induced Guanylate-Binding Proteins: Guardians of Host Defense in Health and Disease.* Journal of Experimental Medicine, 2019. **216**(3): p. 482-500.
8. Xavier, A., et al., *Hgbp1 Coordinates Chlamydia Restriction and Inflammasome Activation through Sequential Gtp Hydrolysis.* Cell Reports, 2020. **31**(7).
9. Fisch, D., et al., *Human Gbp 1 Is a Microbe-Specific Gatekeeper of Macrophage Apoptosis and Pyroptosis.* The EMBO Journal, 2019. **38**(13).
10. Tripal, P., et al., *Unique Features of Different Members of the Human Guanylate-Binding Protein Family.* Journal of Interferon & Cytokine Research, 2007. **27**(1): p. 44-52.
11. Shydlovskyi, S., et al., *Nucleotide-Dependent Farnesyl Switch Orchestrates Polymerization and Membrane Binding of Human Guanylate-Binding Protein 1.* Proceedings of the National Academy of Sciences, 2017. **114**(28).
12. Prakash, B., et al., *Structure of Human Guanylate-Binding Protein 1 Representing a Unique Class of Gtp-Binding Proteins.* Nature, 2000. **403**(6769): p. 567-571.
13. Ghosh, P., et al., *Structural Mechanism for Guanylate-Binding Proteins (Gbps) Targeting by the Shigella E3 Ligase Ipah9.8.* PLOS Pathogens, 2019. **15**(6).
14. Lorenz, C., et al., *Farnesylation of Human Guanylate-Binding Protein 1 as Safety Mechanism Preventing Structural Rearrangements and Uninduced Dimerization.* The FEBS Journal, 2019. **287**(3): p. 496-514.
15. Agam, G., et al., 2022.
16. Kuhnemuth, R. and C.A.M. Seidel, *Principles of Single Molecule Multiparameter Fluorescence Spectroscopy.* Single Molecules, 2001. **2**(4): p. 251-254.

17. Hendrix, J. and D.C. Lamb, *Pulsed Interleaved Excitation: Principles and Applications.* Fluorescence Fluctuation Spectroscopy (Ffs), Pt A, 2013. **518**: p. 205-243.
18. Peulen, T.-O., et al. *Integrative Dynamic Structural Biology Unveils Conformers Essential for the Oligomerization of a Large Gtpase*. 2020. arXiv:2004.04229.
19. Dimura, M., et al., *Automated and Optimally Fret-Assisted Structural Modeling.* Nature Communications, 2020. **11**(1).
20. Ahmed, A. and H. Gohlke, *Multiscale Modeling of Macromolecular Conformational Changes Combining Concepts from Rigidity and Elastic Network Theory.* Proteins-Structure Function and Bioinformatics, 2006. **63**(4): p. 1038-1051.
21. Peulen, T.O., O. Opanasyuk, and C.A.M. Seidel, *Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved Fret Measurements of Labeled Macromolecules Accurately.* Journal of Physical Chemistry B, 2017. **121**(35): p. 8211-8241.
22. Barth, A., et al., *Unraveling Multi-State Molecular Dynamics in Single-Molecule Fret Experiments. I. Theory of Fret-Lines.* Journal of Chemical Physics, 2022. **156**(14).
23. Felekyan, S., et al., *Filtered Fcs: Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules.* Chemphyschem, 2012. **13**(4): p. 1036-1053.
24. Vopel, T., et al., *Triphosphate Induced Dimerization of Human Guanylate Binding Protein 1 Involves Association of the C-Terminal Helices: A Joint Double Electron-Electron Resonance and Fret Study.* Biochemistry, 2014. **53**(28): p. 4590-4600.
25. Cui, W., et al., *Structural Basis for Gtp-Induced Dimerization and Antiviral Function of Guanylate-Binding Proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2021. **118**(15).
26. Sisamakis, E., et al., *Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection.* Methods in Enzymology, Vol 475: Single Molecule Tools, Pt B, 2010. **475**: p. 455-514.
27. Shydlovskyi, S., et al., *Nucleotide-Dependent Farnesyl Switch Orchestrates Polymerization and Membrane Binding of Human Guanylate-Binding Protein 1.* Proceedings of the National Academy of Sciences of the United States of America, 2017. **114**(28): p. E5559-E5568.
28. Sistemich, L., et al., *Structural Requirements for Membrane Binding of Human Guanylate-Binding Protein 1.* The FEBS Journal, 2021. **288**(13): p. 4098-4114.
29. Budde, J.-H., et al. *Fret Nanoscopy Enables Seamless Imaging of Molecular Assemblies with Sub-Nanometer Resolution*. 2021. arXiv:2108.00024.
30. Sistemich, L., et al., *The Molecular Mechanism of Polymer Formation of Farnesylated Human Guanylate-Binding Protein 1.* Journal of Molecular Biology, 2020. **432**(7): p. 2164-2185.

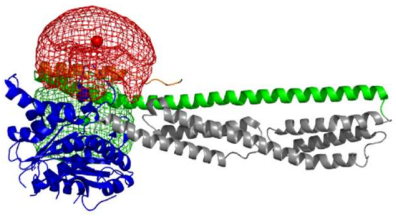
31. Sistemich, L. and C. Herrmann, *Purification of and Characterization of Its Polymerization and Membrane Binding*, in *Dynamin Superfamily Gtpases*. 2020. p. 67-81.
32. Ince, S., et al., *The Human Guanylate-Binding Proteins Hgbp-1 and Hgbp-5 Cycle between Monomers and Dimers Only.* The FEBS Journal, 2017. **284**(14): p. 2284-2301.
33. Widengren, J., et al., *Single-Molecule Detection and Identification of Multiple Species by Multiparameter Fluorescence Detection.* Analytical Chemistry, 2006. **78**(6): p. 2039-2050.
34. Kudryavtsev, V., et al., *Combining Mfd and Pie for Accurate Single-Pair Forster Resonance Energy Transfer Measurements.* Chemphyschem, 2012. **13**(4): p. 1060-1078.
35. Hellenkamp, B., et al., *Precision and Accuracy of Single-Molecule Fret Measurements-a Multi-Laboratory Benchmark Study (Vol 15, Pg 984, 2018).* Nature Methods, 2018. **15**(11): p. 984-984.
36. Sindbert, S., et al., *Accurate Distance Determination of Nucleic Acids Via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity.* Journal of the American Chemical Society, 2011. **133**(8): p. 2463-2480.
37. Dimura, M., et al., *Quantitative Fret Studies and Integrative Modeling Unravel the Structure and Dynamics of Biomolecular Systems.* Current Opinion in Structural Biology, 2016. **40**: p. 163-185.
38. Antonik, M., et al., *Separating Structural Heterogeneities from Stochastic Variations in Fluorescence Resonance Energy Transfer Distributions Via Photon Distribution Analysis.* Journal of Physical Chemistry B, 2006. **110**(13): p. 6970-6978.
39. Kalinin, S., et al., *Detection of Structural Dynamics by Fret: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States.* Journal of Physical Chemistry B, 2010. **114**(23): p. 7983-7995.
40. Wehner, M., S. Kunzelmann, and C. Herrmann, *The Guanine Cap of Human Guanylate-Binding Protein 1 Is Responsible for Dimerization and Self-Activation of Gtp Hydrolysis.* FEBS Journal, 2012. **279**(2): p. 203-210.
41. Hengstenberg, C.S., *Structural Dynamics and Implications for Dimer Formation of Human Guanylate-Binding Protein 1*. 2015.

Supplementary figures

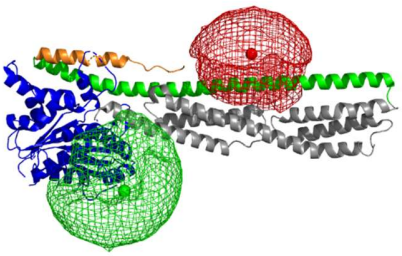
K215C/K567C



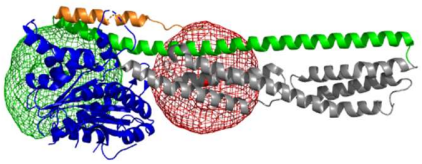
K232C/D578C



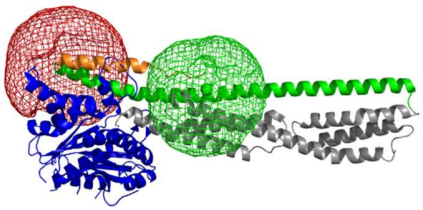
A25C/E514C



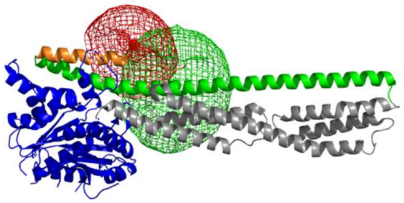
T197C/E366C



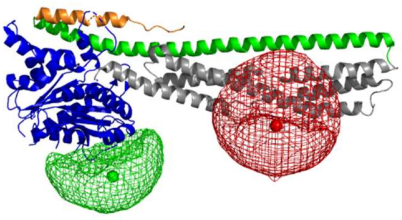
K382C/K562C



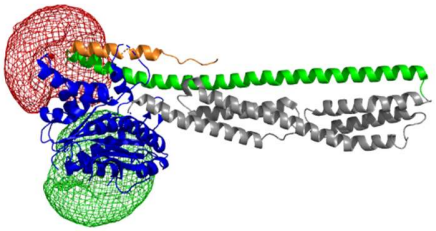
E533C/M583C



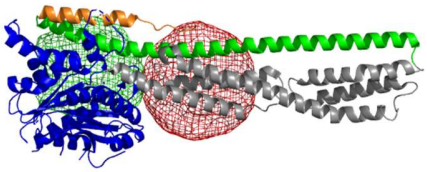
E4C/D405C



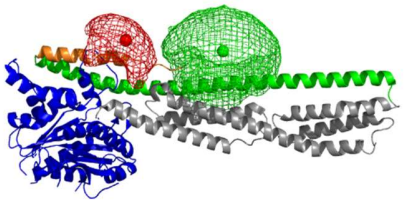
A267C/E563C



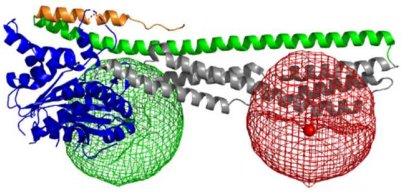
K232C/E366C



E521C/K582C



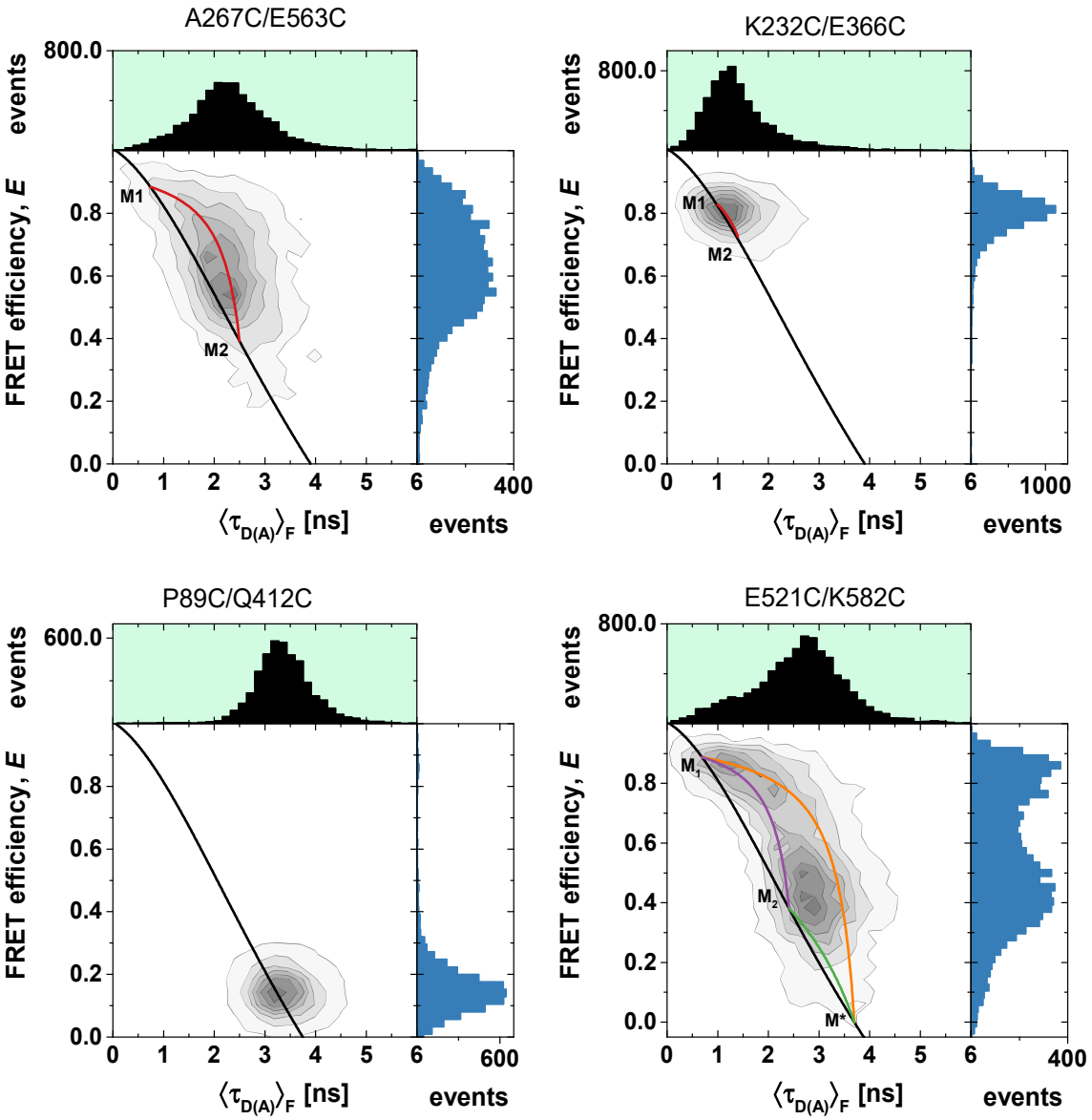
P89C/Q412C

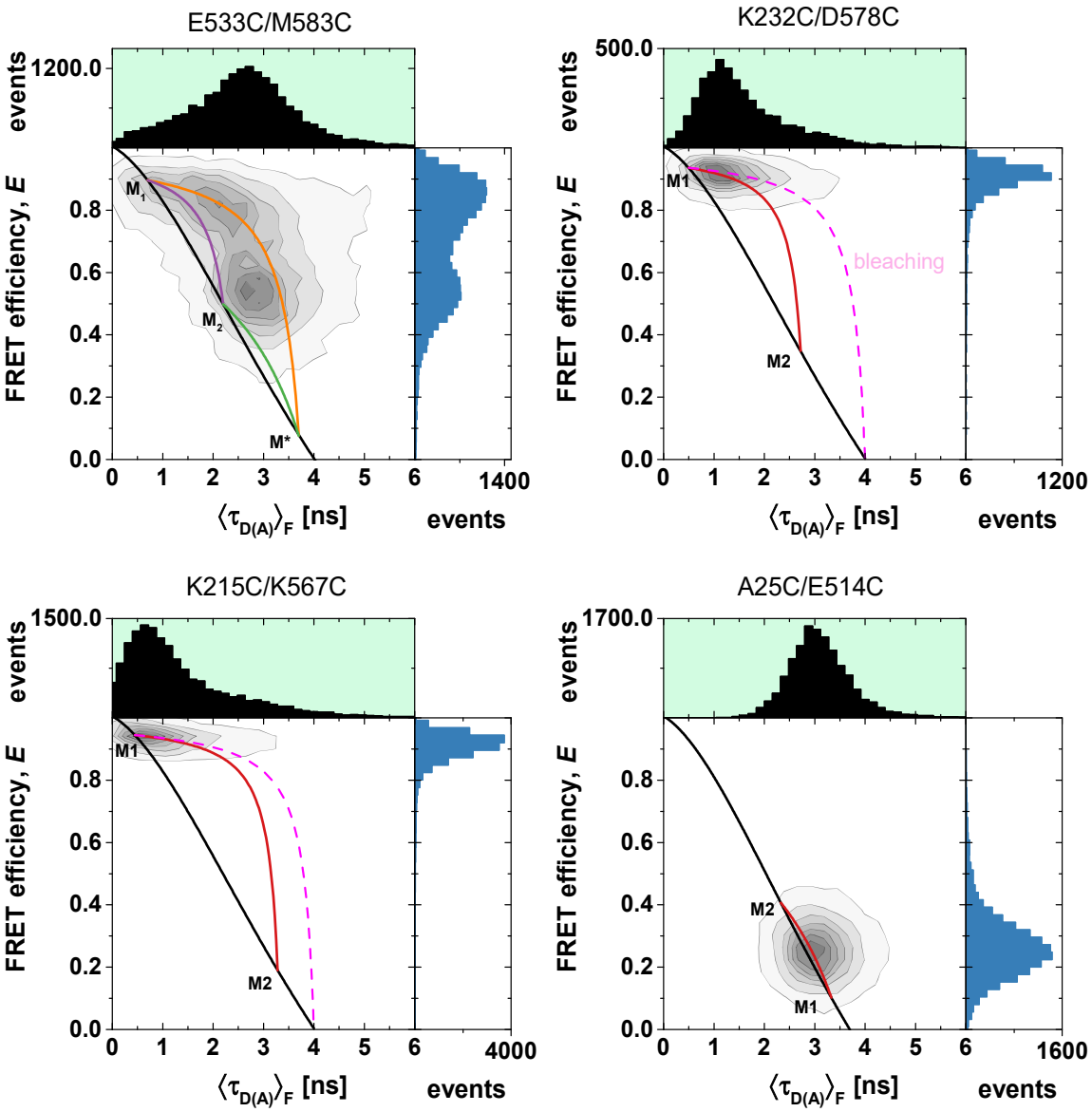


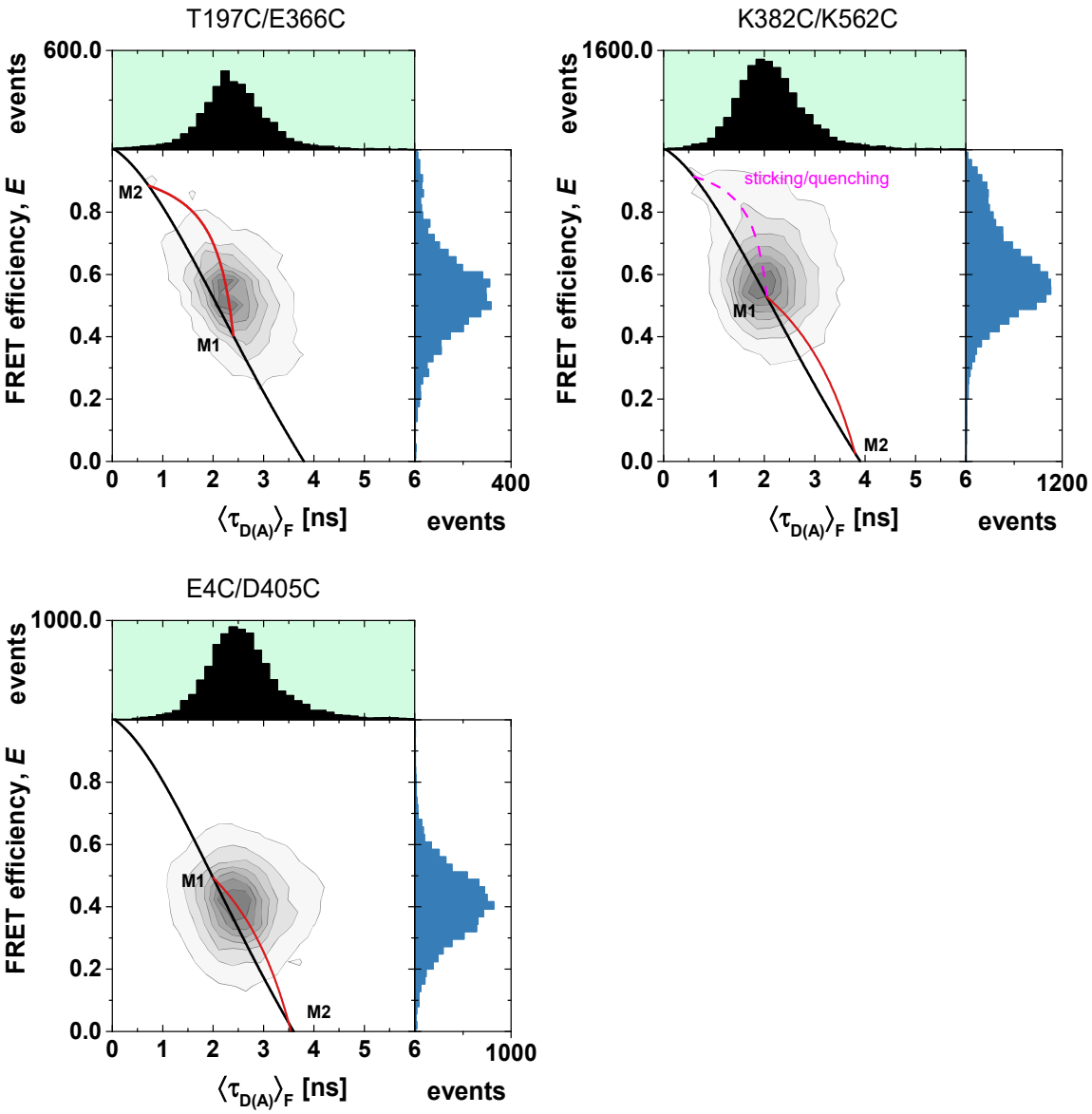
Supplementary Figure 13: 3d structures of non-farnesylated hGBP1 with accessible volumes of dyes for each mutant. Green volume shows the accessible volume of the donor dye (Alexa488), red volume shows the accessible volume of the acceptor dye (Alexa647). Mean position of the dye is indicated via a colored sphere in the accessible volumes.

Supplementary Table 7: Labeling efficiencies of used hGBP1 samples. Labeling efficiencies (LE) determined after labeling of the purified hGBP1 variants, which are shown in their full and abbreviated names. Cys9 means the mutation of all-natural cysteines (C12A, C82A, C225S, C235A, C270A, C311S, C396A, C407S, C589S), while Cys8 possess all these mutations except C589S in order to allow for farnesylation. The labeling efficiencies were determined at the same time as the protein concentration by measuring the absorbance of the sample at 280 nm (ε=45400 $M^{-1}cm^{-1}$), 491 nm (Alexa 488 ε=71000 $M^{-1}cm^{-1}$) and 651 nm (Alexa 647 ε=268000 $M^{-1}cm^{-1}$). Overlap of the absorbances was accounted for and corrected accordingly.

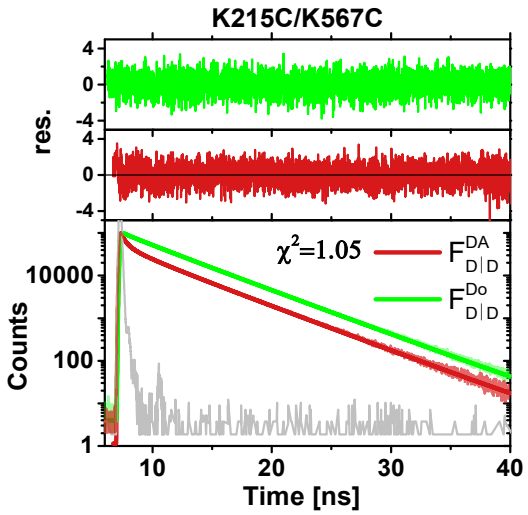
Variant Full Name	Abbreviation	LE Alexa 488 [%]	LE Alexa 647 [%]
hGBP1 Cys9 K215C K567C	Cys9 215-567	103	74
hGBP1 Cys9 K232C D578C	Cys9 232-578	109	73
hGBP1 Cys9 A25C E514C	Cys9 25-514	102	84
hGBP1 Cys9 T197C E366C	Cys9 197-366	100	82
hGBP1 Cys9 K382C K562C	Cys9 382-562	91	92
hGBP1 Cys9 E533C M583C	Cys9 533-583	101	66
hGBP1 Cys9 E4C D405C	Cys9 4-405	80	88
hGBP1 Cys9 A267C E563C	Cys9 267-563	57	85
hGBP1 Cys9 K232C E366C	Cys9 232-366	98	71
hGBP1 Cys9 E521C K582C	Cys9 521-582	86	69
hGBP1 Cys9 P89C Q412C	Cys9 89-412	87	97
hGBP1 Cys8 N18C Q577C	Cys8 18-577	103	89
hGBP1 Cys8 N18C Q344C	Cys8 18-344	91	88
hGBP1 Cys8 Q344C A496C	Cys8 344-496	107	84
hGBP1 Cys8 Q344C Q525C	Cys8 344-525	103	84
hGBP1 Cys8 E533C M583C	Cys8 533-583	87	71

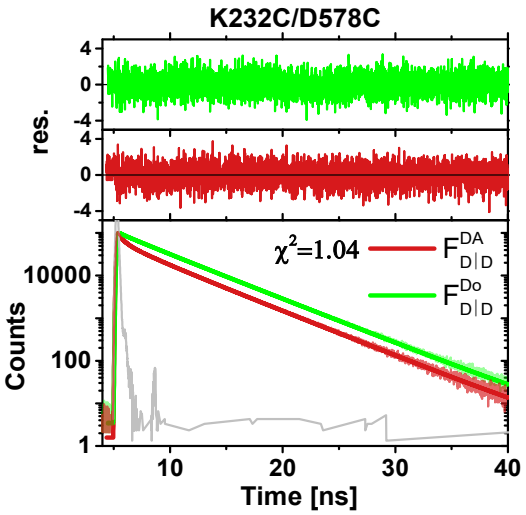

A267C/E563C
K232C/E366C
P89C/Q412C
E521C/K582C
events
FRET efficiency, E
M1
M2
M*
⟨τD(A)⟩F [ns]
events

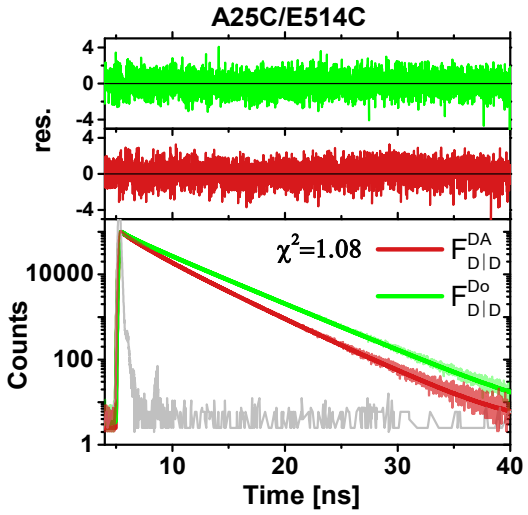

E533C/M583C
events
1200.0
FRET efficiency, E
M1
M2
M*
0.0
0.2
0.4
0.6
0.8
0 1 2 3 4 5 6
⟨τD(A)⟩F [ns]
1400
events
K232C/D578C
500.0
M1
bleaching
M2
1200
K215C/K567C
1500.0
4000
A25C/E514C
1700.0
1600

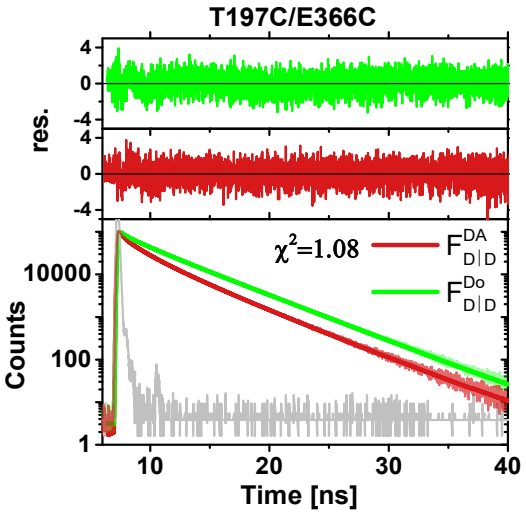


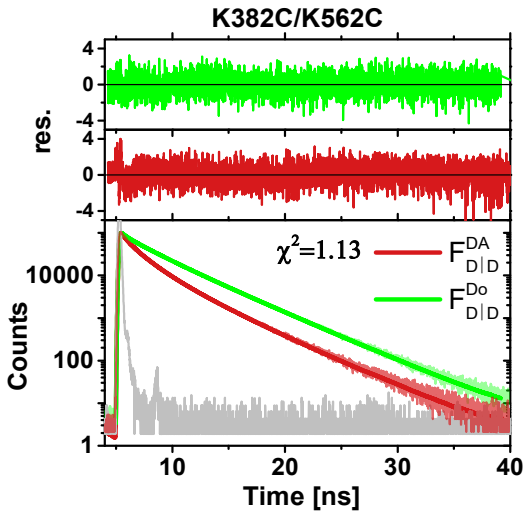
Supplementary Figure 14: Single-molecule fluorescence measurements. 2d-MFD-histograms of the non-farnesylated hGBP1 samples. hGBP1 was labeled using Alexa488 as donor and Alexa647 as acceptor dye. Top histogram shows the distribution of fluorescence weighted donor lifetimes in presence of an acceptor, $\langle \tau_{D(A)} \rangle_f$. Right histogram shows the distribution of FRET efficiencies. Only bursts are shown from double labeled molecules, filtering was done using a stoichiometry and ALEX2CDE cut. Black line represents the static FRET-line, red the dynamic between the states M_1 and M_2, and in more complex case as in E533C/M583C and E521C/K582C dynamic lines are shown in green, purple and orange. Dashed magenta line indicates bleaching behavior which could not be completely filtered out.

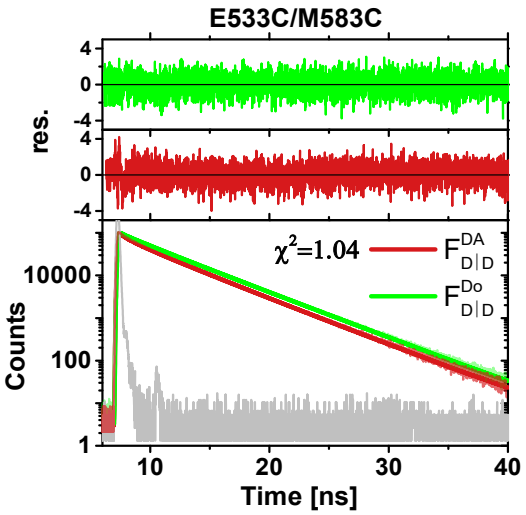

K215C/K567C
res.
Counts
Time [ns]
χ²=1.05
F^DA D|D
F^Do D|D

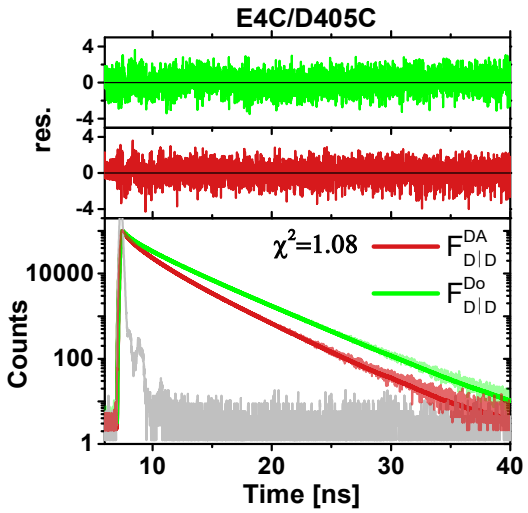

K232C/D578C
res.
Counts
Time [ns]
χ²=1.04
F^DA D|D
F^Do D|D

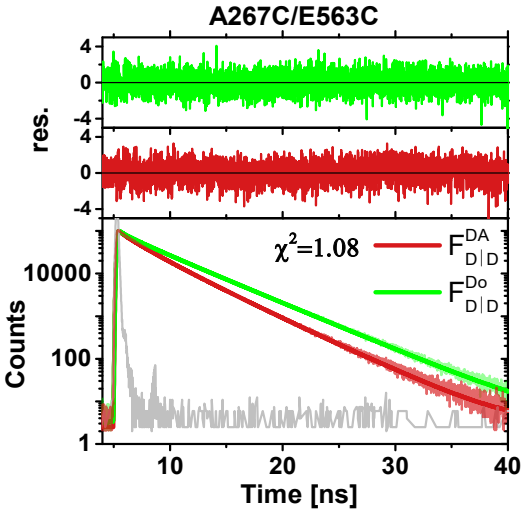

A25C/E514C
res.
Counts
Time [ns]
χ²=1.08
F^DA D|D
F^Do D|D

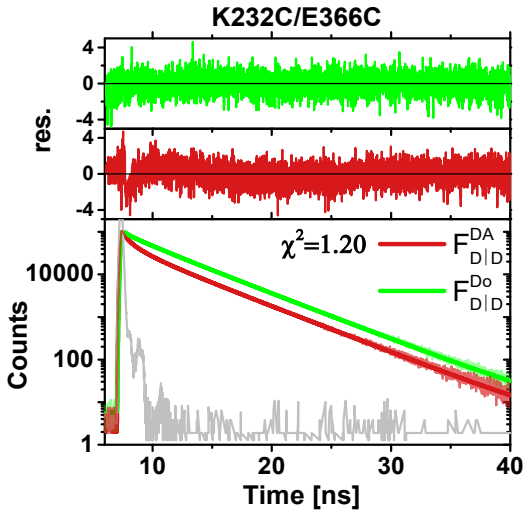

T197C/E366C
res.
Counts
Time [ns]
χ²=1.08
F^DA D|D
F^Do D|D

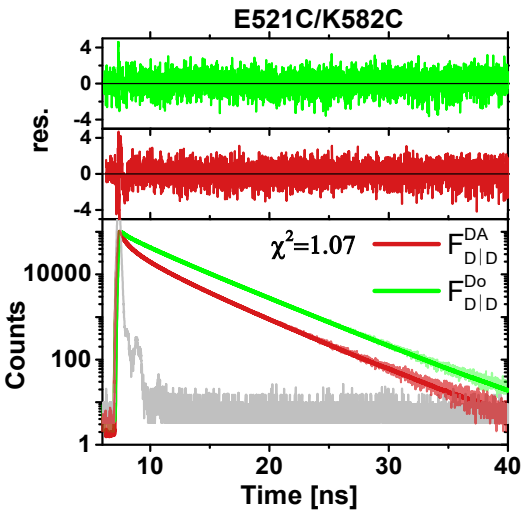

K382C/K562C
res.
Counts
Time [ns]
χ²=1.13
F^DA D|D
F^Do D|D

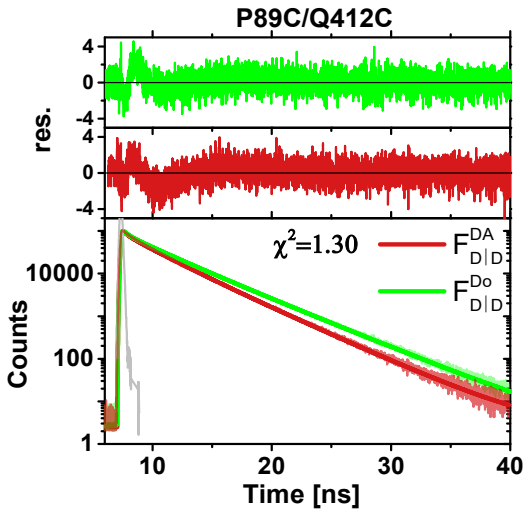

E533C/M583C
res.
Counts
Time [ns]
χ²=1.04
F^DA D|D
F^Do D|D


E4C/D405C
res.
Counts
Time [ns]
χ²=1.08
F^DA_D|D
F^Do_D|D

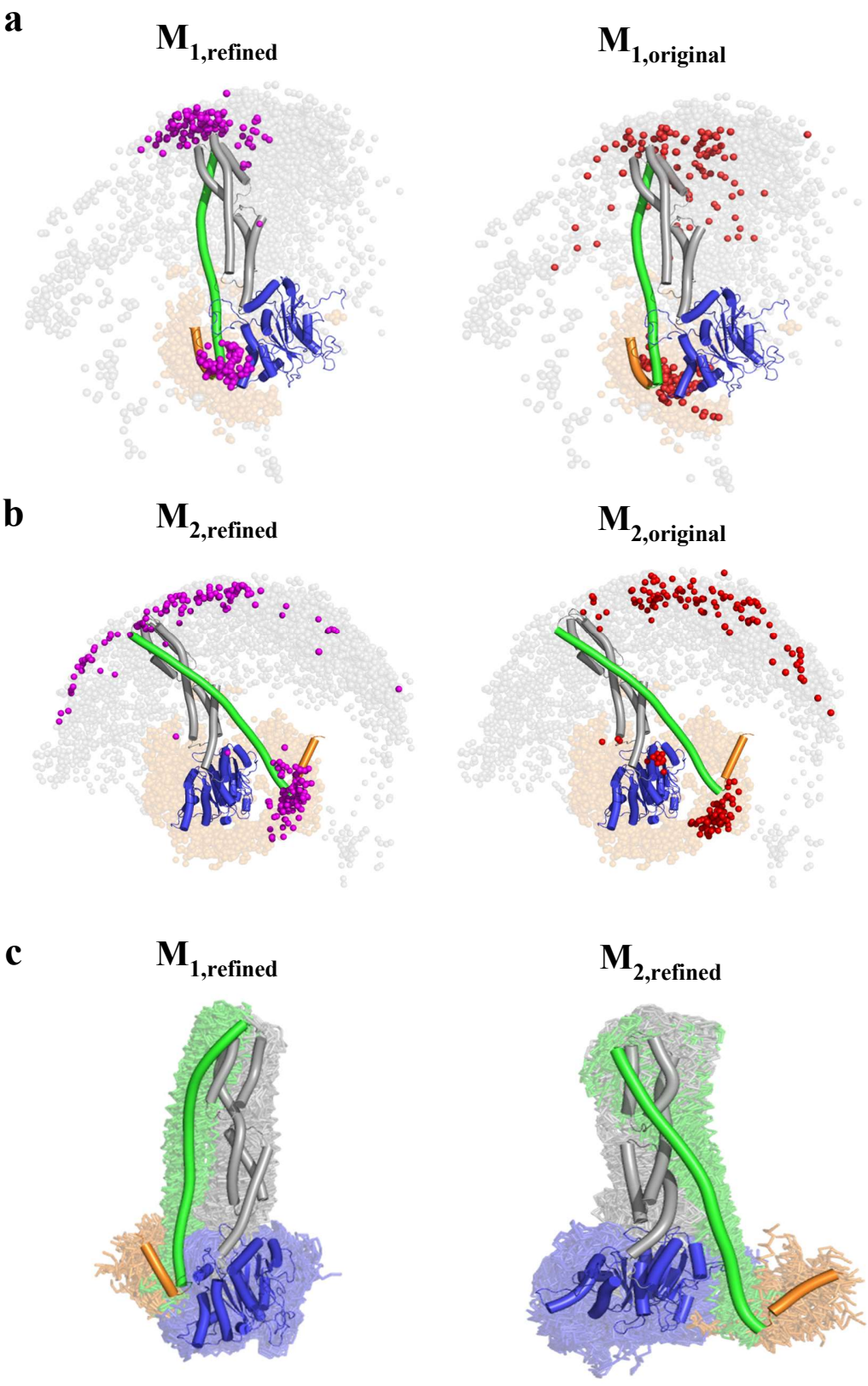

A267C/E563C
res.
Counts
Time [ns]
χ²=1.08
F^DA_D|D
F^Do_D|D


K232C/E366C
res.
Counts
Time [ns]
χ²=1.20
F^DA_D|D
F^Do_D|D

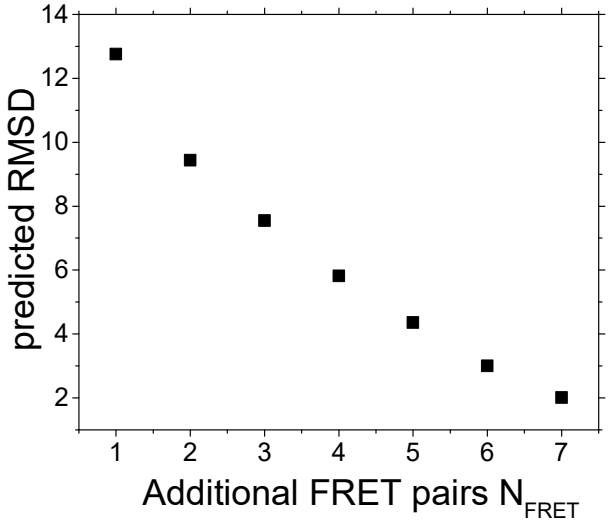

E521C/K582C
res.
Counts
Time [ns]
χ²=1.07
F^DA_D|D
F^Do_D|D



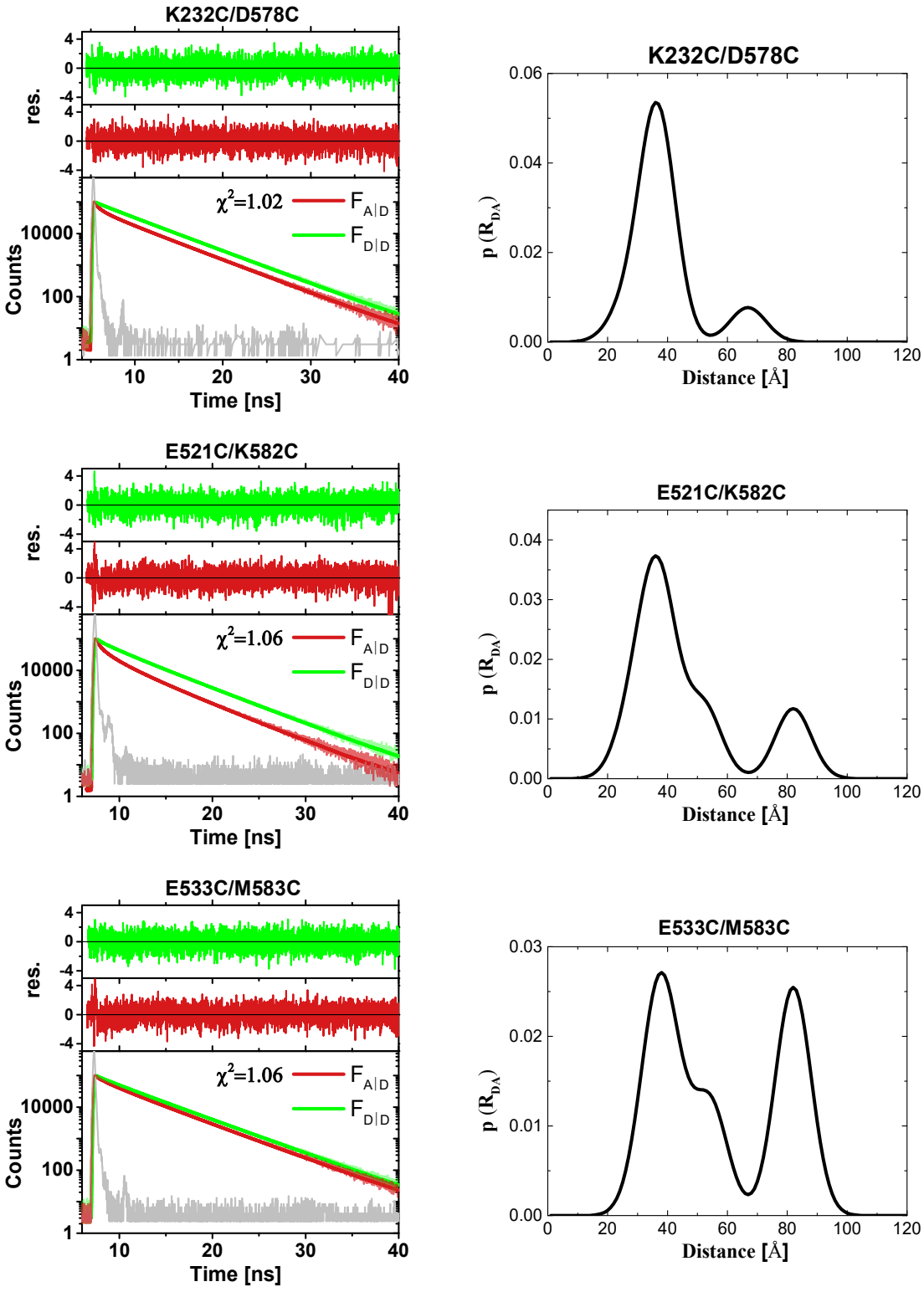
Supplementary Figure 15: Ensemble Lifetime fits. Ensemble fluorescence intensity decays of the donor dye in absence of an acceptor (green, $F_{\mathrm{D|D}}^{\mathrm{Do}}$) and in presence of an acceptor ($F_{\mathrm{D|D}}^{\mathrm{DA}}$). The fluorescence decay were fitted jointly using a global lifetime fit for the donor only decay and a Gaussian distribution model for the FRET active sample. Donor only and donor-acceptor measurements were done separately using only single donor labeled molecules (Alexa488) and double labeled molecules (Alexa488/Alexa647). Grey shows the Instrument Response Function (IRF).


a
$M_{1,refined}$
$M_{1,original}$
b
$M_{2,refined}$
$M_{2,original}$
c
$M_{1,refined}$
$M_{2,refined}$

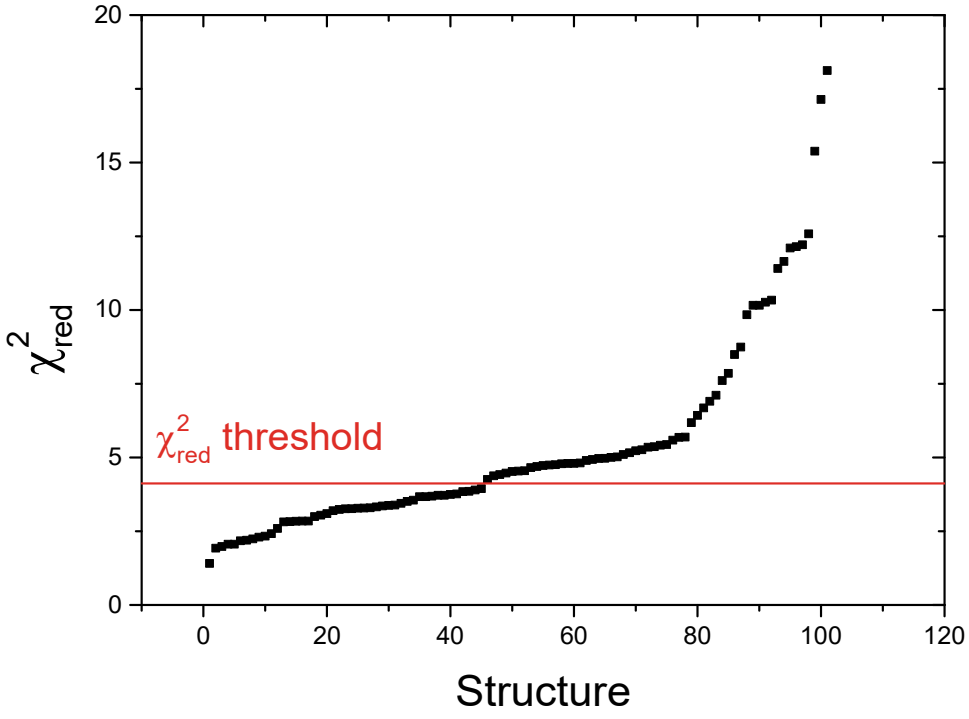
Supplementary Figure 16: Structural refinement of M_1 and M_2 state of non-farnesylated hGBP1. Structural ensemble of hGBP1 monomeric state 1 and 2 using a Meta-Analysis of EPR, SAXS and FRET including N_{FRET}=23. Structures are aligned to the LG-domain. Grey spheres represent the whole ensemble (around 3000 structures) for the position of the MD-domain (amino acid 481), orange spheres the whole ensemble for the position of the beginning point of the α13-helix (amino acid 565), magenta spheres the refined ensemble including N_{FRET}=23, red spheres the original ensemble including N_{FRET}=11. (**a**) Refined M_1 ensemble (left) in comparison to the original M_1 ensemble. (**b**) Refined M_2 ensemble (left) in comparison to the original M_2 ensemble. (**c**) Global alignment of structures with a p-value=0.68 for monomeric state 1 (left) and monomeric state 2 (right). The α12 helix (green) is orientated on the other side of the middle domain for M_2. Additionally the α13 (orange) helix is orientated differently, showing a wider angle in respect to the α12 helix



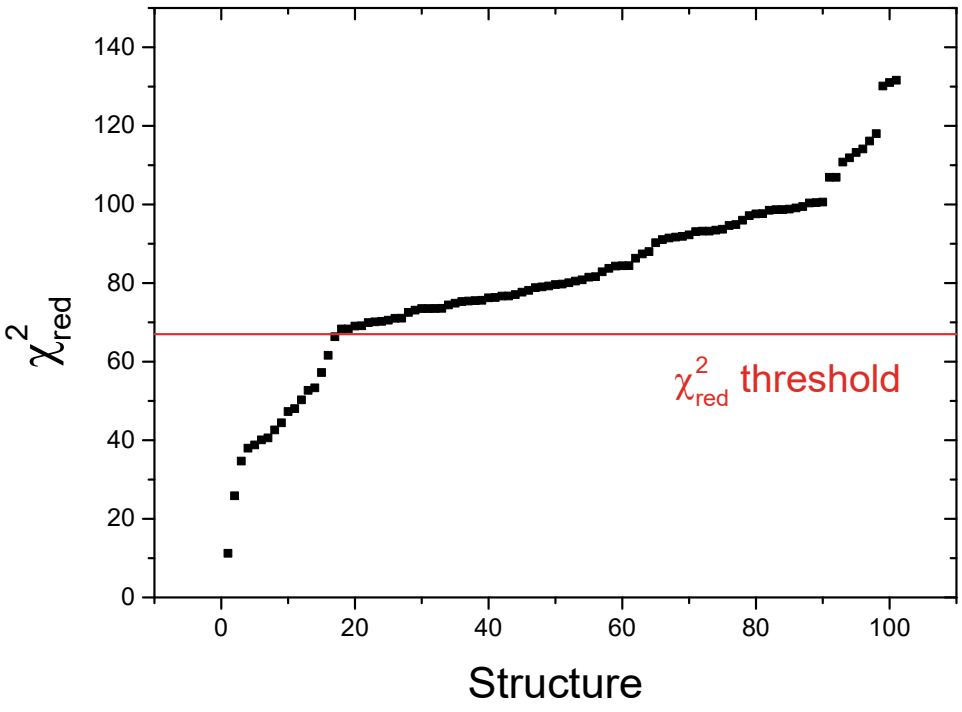
Supplementary Figure 17: Theoretical prediction of RMSD-width of selected ensemble depending on the additional number of FRET pairs included in the analysis [19]. Starting point is the already resolved ensemble [18] using SAXS, EPR and N_{FRET}=11. Additional number of FRET pairs N_{FRET} will narrow down the ensemble.



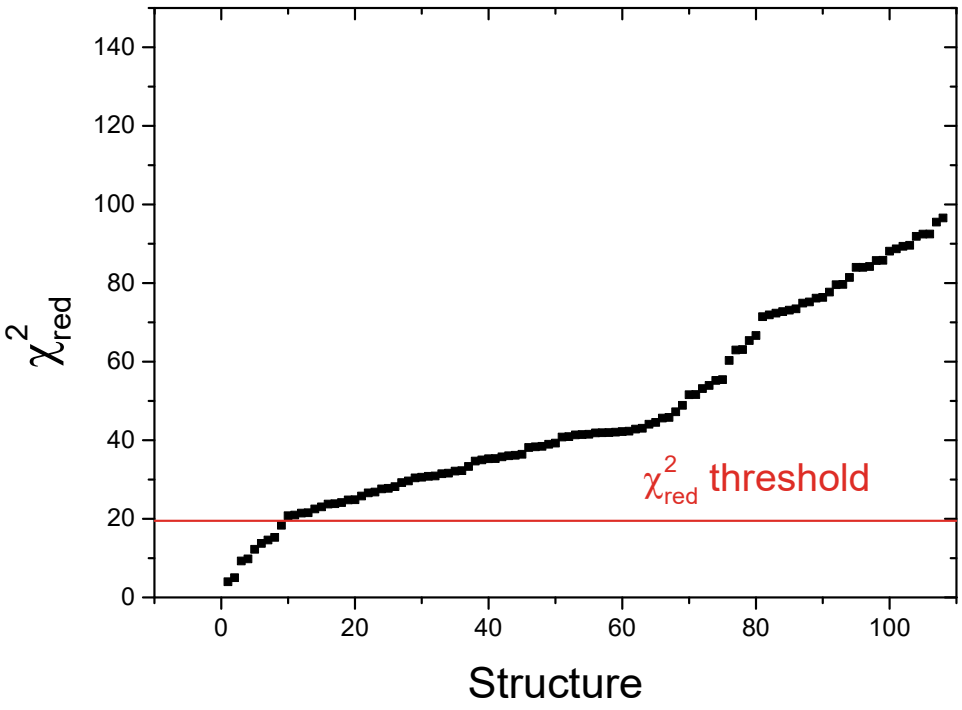
Supplementary Figure 18: Ensemble Lifetime fits. (left) Ensemble fluorescence intensity decays of the donor dye in absence of an acceptor (green, $F_{\mathrm{D|D}}^{\mathrm{Do}}$) and in presence of an acceptor ($F_{\mathrm{D|D}}^{\mathrm{DA}}$) for the mutants where a minor state $M_{1,\,\mathrm{attached}}$ was found. (right) Distance distribution based on three (K232C/D578C) or four (E521C/K582C, E533C/M583C) Gaussian distributions.



Supplementary Figure 19: Selection of $M_{1,attached}$ ensemble. χ^2_{red} of $M_{1,\alpha13\ near}$ ensemble using the additional minor distances found in three mutants (see **Supplementary Figure 18** and **Supplementary Table 10**). Structures above the threshold (red line) were rejected.

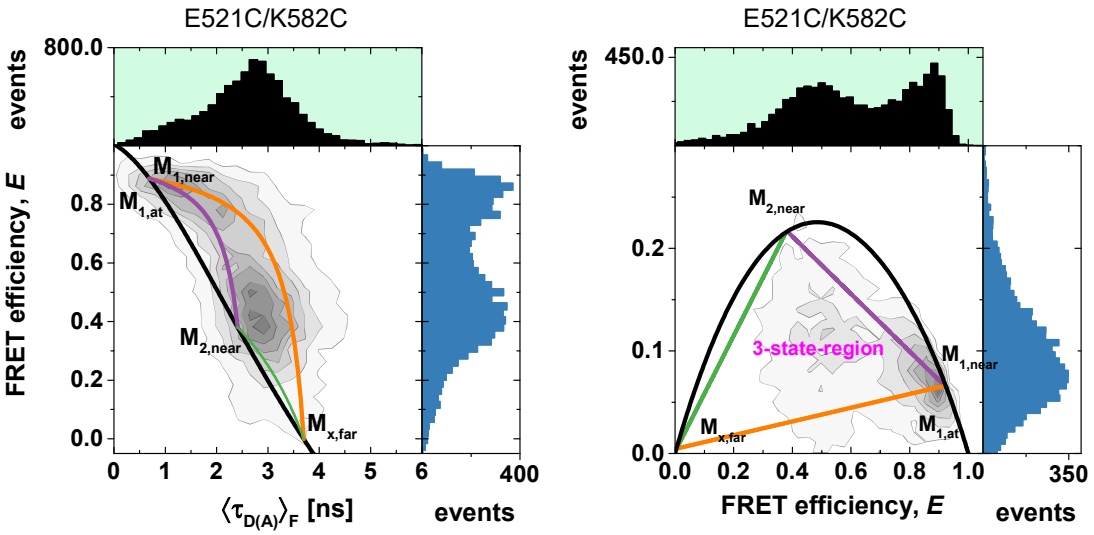


Supplementary Figure 20: Selection of $M_{1,\alpha13\ far}$ ensemble. χ^2_{red} of $M_{1,\alpha13\ near}$ ensemble using the additional minor distances found in three mutants (see **Supplementary Figure 18** and **Supplementary Table 10**). Structures above the threshold (red line) were rejected.



Supplementary Figure 21: Selection of $M_{2,\alpha13\ far}$ ensemble. χ^2_{red} of $M_{2,\alpha13\ near}$ ensemble using the additional minor distances found in three mutants (see **Supplementary Figure 18** and

Supplementary Table 10Supplementary Table 10). Structures above the threshold (red line) were rejected.



Supplementary Figure 22: 2D-sm-MFD histogram of E521C/K582C with the distribution of lifetimes of the donor in presence of an acceptor (top, black histogram) and the distribution of FRET efficiency values (right, blue distribution). Black line indicates the static FRET line, orange the dynamic between $M_{1\ \alpha13\ near}$ and $M_{1,\alpha13\ far}$, purple between $M_{1,\alpha13\ near}$ and $M_{2,\alpha13\ near}$, green between $M_{2,\alpha13\ near}$ and $M_{2,\alpha13\ far}$.

Supplementary Table 8: Overview of measurements for the non-farnesylated hGBP1 mutants based on ensemble TCSPC shown in **Supplementary Figure 15**. Shown are resolved distances R_{DA} for $M_{1,\alpha13\ near}$ and $M_{2,\alpha13\ near}$ state (for additional distances see **Supplementary Table 10**) based on a Gaussian distribution model with width σ. *R*- and *R*+ represent total error for each distance. $r_{\infty,D}$ and $r_{\infty,A}$ represent the residual anisotropy of the donor and acceptor obtained in the sm-MFD measurement.

Mutant	$R_{DA}(M_{1,\alpha13\ near})$ [Å]	σ [Å]	error *R*- [Å]	error *R*+ [Å]	$R_{DA}(M_{2,\alpha13\ near})$ [Å]	σ [Å]	error *R*+ [Å]	error *R*- [Å]	$r_{\infty,D}$	$r_{\infty,A}$
K215C/K567C	36	6	4	5	67	6	27	23	0.19	0.26
K232C/D578C	38	6	5	6	59	6	24	21	0.20	0.36
A25C/E514C	68	6	6	7	55	6	10	10	0.21	0.28
T197C/E366C	59	6	5	5	39	6	4	4	0.18	0.27
K382C/K562C	54	6	6	6	82	6	6	5	0.22	0.30
E533C/M583C	39	6	3	4	54	6	14	11	0.19	0.29
E4C/D405C	50	6	5	6	73	6	9	9	0.26	0.29
A267C/E563C	38	6	5	5	58	6	6	6	0.30	0.27
K232C/E366C	40	6	5	5	44	6	5	5	0.18	0.34
E521C/K582C	37	6	3	4	52	6	5	8	0.13	0.29
P89C/Q412C	71	6	6	6	68	6	6	6	0.28	0.19

Supplementary Table 9: Overview of the calculated errors. Errors consist out of a statistical error of the data, error due to the dye simulation as accessible volumes on the molecule, reference uncertainty of the donor only sample, and uncertainty of the orientation factor of the dye.

		total		statistical error		Dye simulation		Reference uncertainty		Orientation factor	
Mutant	**state**	***R*- [Å]**	***R*+ [Å]**	**Δ-**	**Δ+**	**Δ-**	**Δ+**	**Δ-**	**Δ+**	**Δ-**	**Δ+**
K215C/K567C	1	4	5	0.036	0.061	0.028	0.057	0.001	0.001	0.028	0.057
K232C/D578C	1	5	6	0.016	0.029	0.026	0.054	0.001	0.001	0.026	0.054
A25C/E514C	1	6	7	0.025	0.045	0.015	0.030	0.036	0.029	0.015	0.030
T197C/E366C	1	5	5	0.011	0.027	0.017	0.034	0.014	0.013	0.017	0.034
K382C/K562C	1	6	6	0.022	0.038	0.019	0.038	0.008	0.008	0.019	0.038
E533C/M583C	1	3	4	0.022	0.056	0.026	0.053	0.001	0.001	0.026	0.053
E4C/D405C	1	5	6	0.030	0.057	0.020	0.041	0.005	0.005	0.020	0.041
A267C/E563C	1	5	5	0.016	0.026	0.018	0.036	0.012	0.011	0.018	0.036
K232C/E366C	1	5	5	0.016	0.027	0.025	0.051	0.001	0.001	0.025	0.051
E521C/K582C	1	3	4	0.028	0.059	0.027	0.056	0.001	0.001	0.027	0.056
P89C/Q412C	1	6	6	0.007	0.015	0.014	0.029	0.049	0.036	0.014	0.029
K215C/K567C	2	17	18	0.126	0.249	0.015	0.015	0.200	0.027	0.015	0.015
K232C/D578C	2	8	9	0.034	0.067	0.017	0.017	0.014	0.013	0.017	0.017
A25C/E514C	2	5	5	0.030	0.019	0.018	0.018	0.009	0.009	0.018	0.018
T197C/E366C	2	3	4	0.032	0.044	0.025	0.025	0.001	0.001	0.025	0.025
K382C/K562C	2	16	10	0.018	0.016	0.012	0.012	0.161	0.075	0.012	0.012
E533C/M583C	2	5	5	0.040	0.015	0.018	0.018	0.008	0.008	0.018	0.018
E4C/D405C	2	8	8	0.028	0.026	0.014	0.014	0.057	0.041	0.014	0.014
A267C/E563C	2	7	7	0.014	0.009	0.017	0.017	0.013	0.012	0.017	0.017
K232C/E366C	2	5	5	0.034	0.039	0.023	0.023	0.002	0.002	0.023	0.023
E521C/K582C	2	4	4	0.027	0.032	0.019	0.019	0.007	0.006	0.019	0.019
P89C/Q412C	2	6	5	0.007	0.007	0.015	0.015	0.038	0.030	0.015	0.015

Supplementary Table 10: Overview of the distances and fractions (below, in %) where additional distances where found, using a fixed width in a Gaussian distribution model of a fluorescence decay using sub-ensemble data of a sm-MFD measurement using PIE selection as described in the methods.

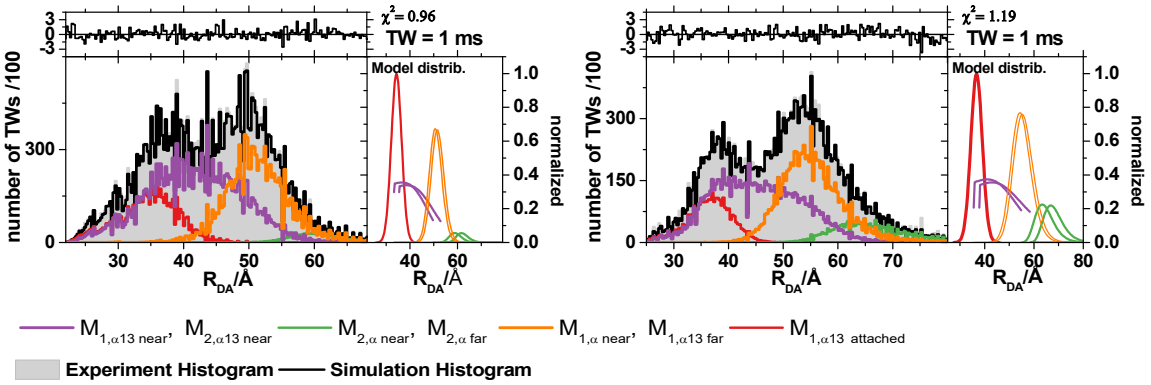
Mutant	Static analysis (σ = 6 Å)			
	$R_{DA}(M_{1,\alpha13\ attached})$ [Å]	$R_{DA}(M_{1,\alpha13\ near})$ [Å]	$R_{DA}(M_{2,\alpha13\ near})$ [Å]	$R_{DA}(M_{x,\alpha13\ far})$ [Å]
E533C/M583C	35 (15%)	39 (29%)	54 (19%)	82 (37%)
E521C/K582C	29 (15%)	37 (49%)	52 (19%)	82 (17%)
K232C/D578C	26 (10%)	37 (78%)	69 (12%)	-

Supplementary Table 11: Simulation of interdye distances for the expected species $M_{1,\alpha13\ near}$, $M_{2,\alpha13\ near}$ and $M_{1,\alpha13\ far}$,$M_{2,\alpha13\ far}$ based on accessible volumes, crystal structure 6k1z, and a modified version of it with a wide angle between the α12/13 helices (around 180°, see **Figure 44**).

	Simulation			
Mutant	$R_{DA}(M_{1,\alpha13\ near})$ [Å]	$R_{DA}(M_{2,\alpha13\ near})$ [Å]	$R_{DA}(M_{1,\alpha13\ far})$ [Å]	$R_{DA}(M_{2,\alpha13\ far})$ [Å]
E533C/M583C	40	51	85	96
E521C/K582C	43	58	71	68
N18C/V577C	64	45	62	49
K232C/D578C	41	61	55	87

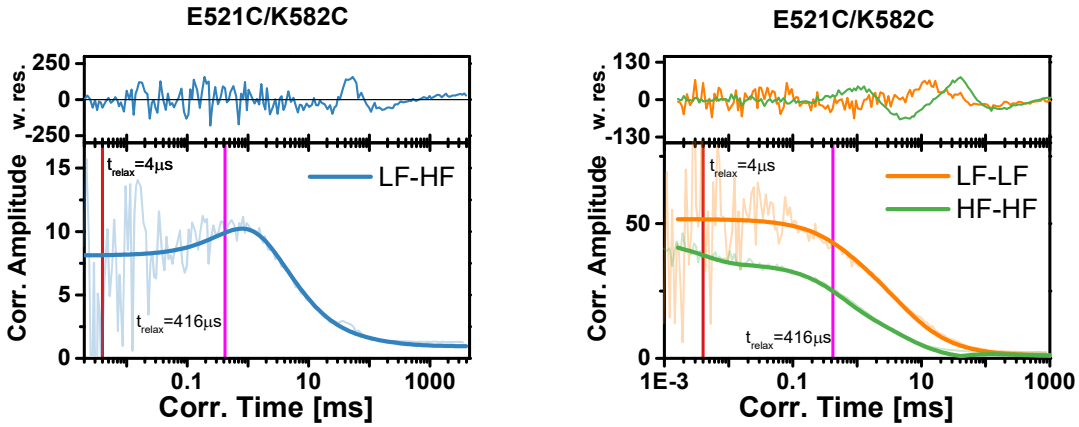
Supplementary Table 12: Resulting fractions for the E533C/M583C and E521C/K582C samples using a linear model with one static and 3 dynamic states and resulting relaxation times of the dynamic states.

	Fraction [%]				Relaxation times [µs]		
Species	$M_{1,\alpha13\ at}$	$M_{1,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ near}$	$M_{1,\alpha13\ far}$ ⇋ $M_{1,\alpha13\ near}$	$M_{2,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ far}$	$M_{1,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ near}$	$M_{1,\alpha13\ far}$ ⇋ $M_{1,\alpha13\ near}$	$M_{2,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ far}$
E533C/M583C	19.1	47.3	31.2	2.5	434	<2	<3
E521C/K582C	13.4	41.7	35.2	9.6	410	<3	<5

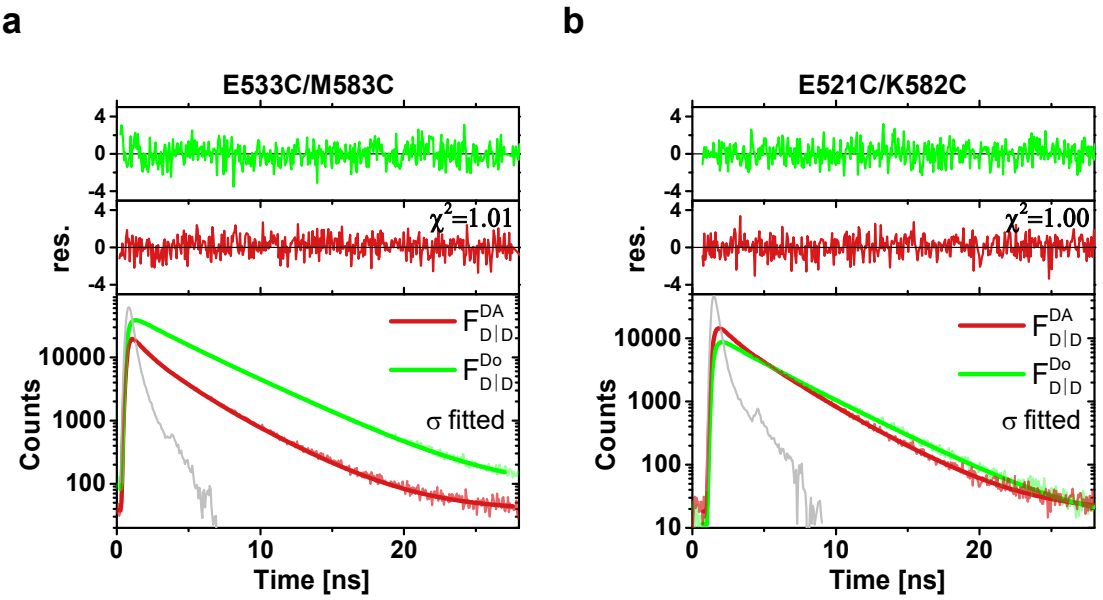


Supplementary Figure 23: PDA of E533C/M583C and E521C/K582C. Static fraction of M_1 is shown in red. Left layer shows the distribution of TW measured (black) and analyzed by the model. Red layer shows the corresponding distribution of the model. Data was analyzed globally using TWs of the size 1, 2 and 3 ms.

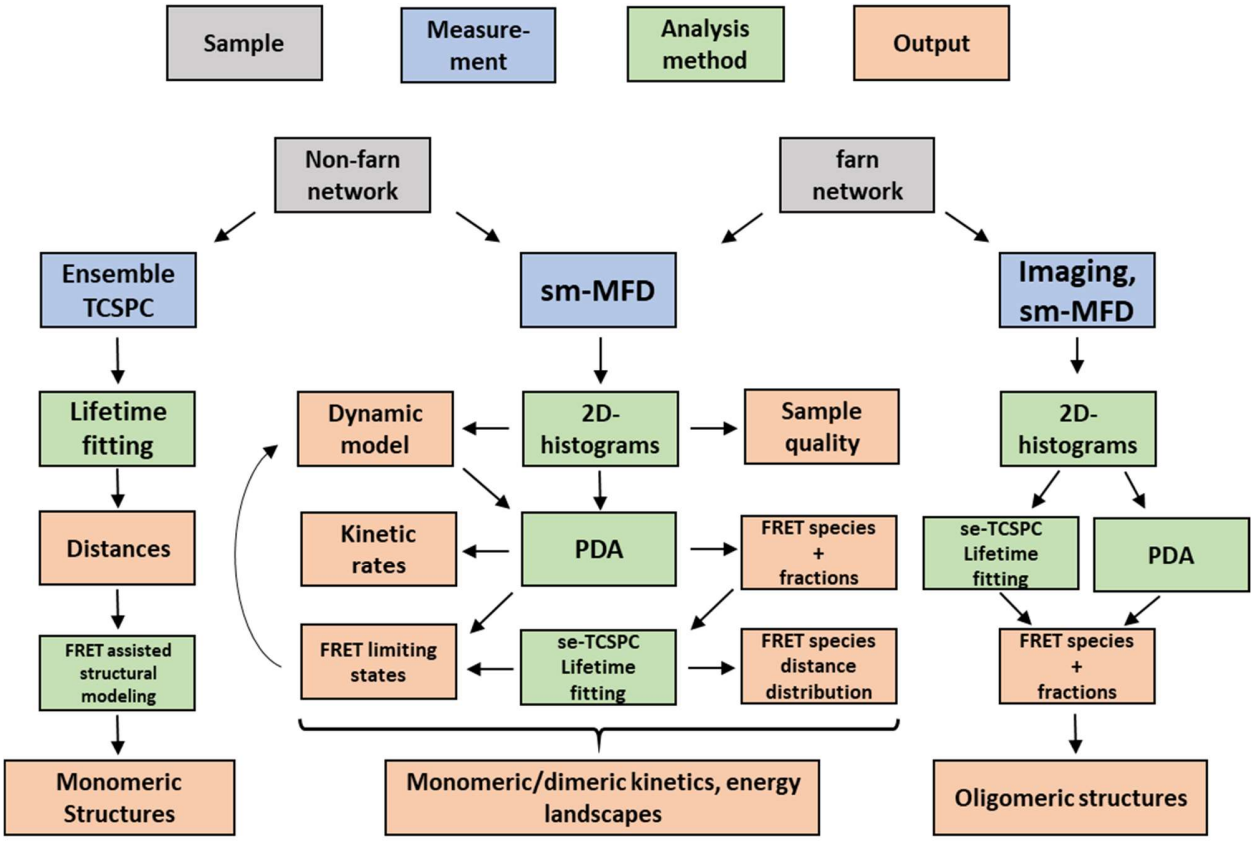
a **b**



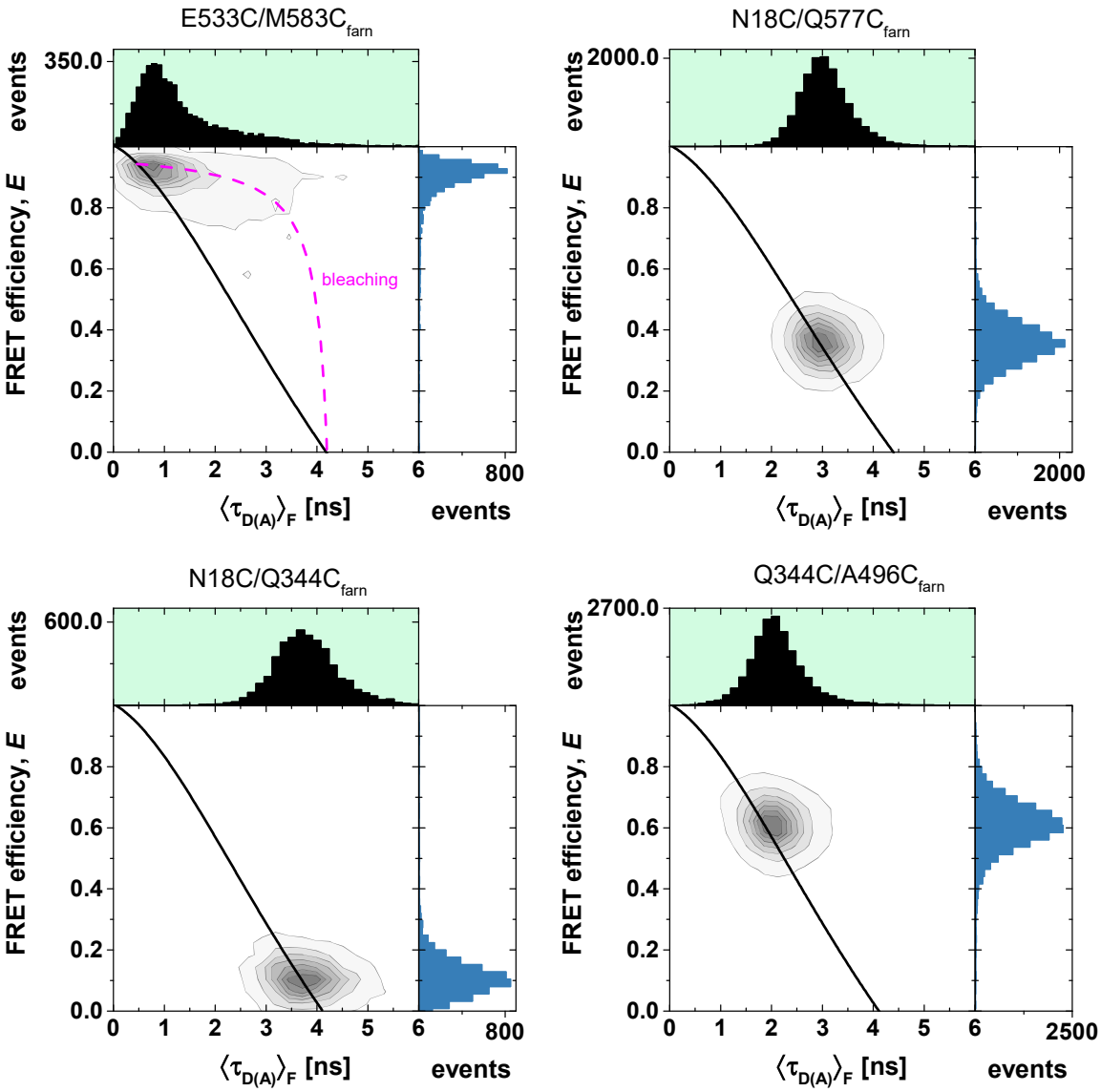
Supplementary Figure 24: fFCS curves and fit for E521C/K582C. a Low FRET (LF) to High FRET (HF) cross correlation curve (blue). Lifetime filters were created via burst selection of only HF and LF molecules. Top graph shows weighted residuals. Red and magenta straight-line show bunching terms with indicated relaxation times responsible for molecular dynamics. b LF-LF (orange) and HF-HF (green) autocorrelation. Cross- and autocorrelation curves were fitted globally.

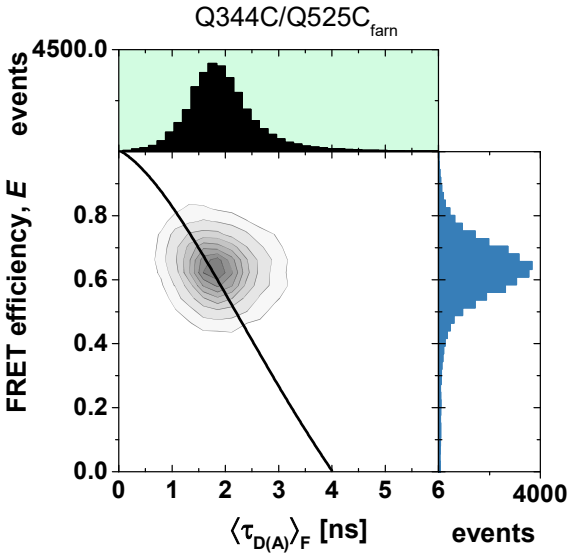


Supplementary Figure 25: Sub-ensemble fuorescence intensity decays of the donor dye in absence of an acceptor (green, $F_{\mathrm{D|D}}^{\mathrm{Do}}$) and in presence of an acceptor ($F_{\mathrm{D|D}}^{\mathrm{DA}}$) for E533C/M583C and E521C/K582C -samples using a fitted Gaussian distribution width σ. The fluorescence decays were fitted jointly using a global lifetime fit for the donor only decay and a Gaussian distribution model for the FRET active sample. Donor only and donor-acceptor measurements were done separately using only single donor labeled molecules (Alexa488) and double labeled molecules (Alexa488/Alexa647). Grey shows the Instrument Response Function (IRF).



Supplementary Figure 26: Overview of the used measurements (blue), the analysis method (green) and output parameter (orange). Arrows indicate connection between models and input/output parameters.


E533C/M583C$_{farn}$
N18C/Q577C$_{farn}$
N18C/Q344C$_{farn}$
Q344C/A496C$_{farn}$
events
FRET efficiency, E
bleaching
$\langle\tau_{D(A)}\rangle_F$ [ns]
events



Supplementary Figure 27: 2D-sm-MFD histograms of all farnesylated hGBP1 samples with the distribution of lifetimes of the donor in presence of an acceptor (top, black histogram) and the distribution of FRET efficiency values (right, blue distribution). Black line indicates the static FRET line, dashed magenta line indicates bleaching of the molecule, that could not be filtered out.

Supplementary Table 13: Overview of resolved distances using seTCSPC and PDA for the sm-measurements of the farnesylated hGBP1 samples in comparison with the predicted distances from the crystal structure (pdb: 6k1z) and the representative structures M_1 and M_2 inferred from the non-farnesylated hGBP1 measurements. Fractions are indicated as %-values in brackets.

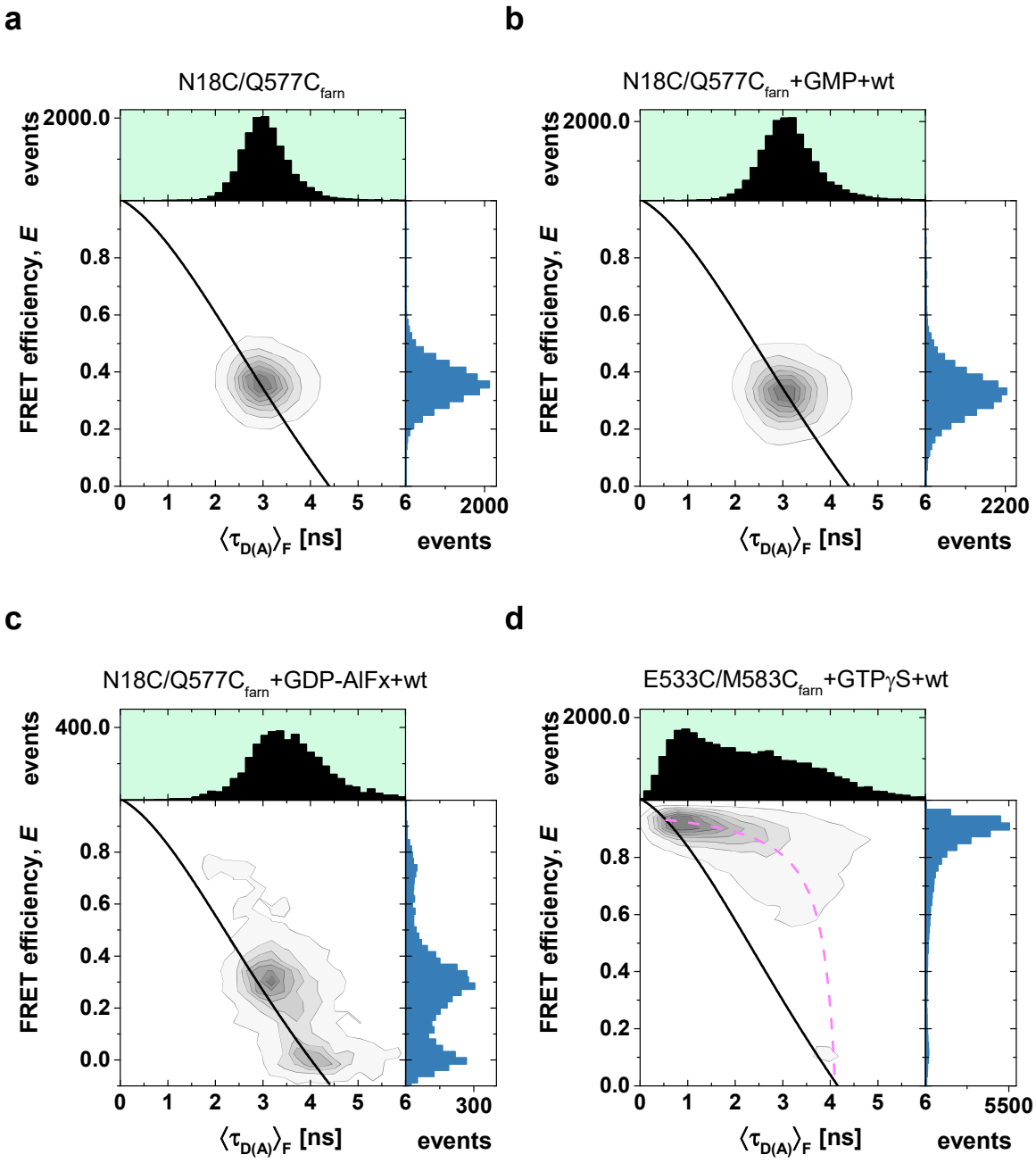
	seTCSPC		PDA			Simulation			Anisotropy	
Sample	R_{DA} [Å]	σ [Å]	R_{DA} [Å]	R_{DA} [Å]	dynamic ratio	6k1z [Å]	M_1 [Å]	M_2 [Å]	ρ_1 (x_1)	ρ_2 (x_2)
N18C/Q344C$_{farn}$	75	6	75		-	66	68	63	0.2ns (60%)	10ns (40%)
N18C/V577C$_{farn}$	58	6	58		-	58.5	66	47	0.2ns (54%)	10ns (46%)
E533C/M583C$_{farn}$	33	6	33 (95%)	55 (1%)	4% (bleaching)	31	37	55	0.2ns (29%)	10ns (71%)
Q344C/A496C$_{farn}$	47	10	47		-	50	48	48	0.2ns (64%)	10ns (36%)
Q344C/Q525C$_{farn}$	48	10	47		-	44	40	43	0.2ns (64%)	10ns (36%)

Supplementary Table 14: Comparison of the distances measured using farnesylated hGBP1 to the crystal structure (pdb: 6k1z) and the representative structures M_1 and M_2.

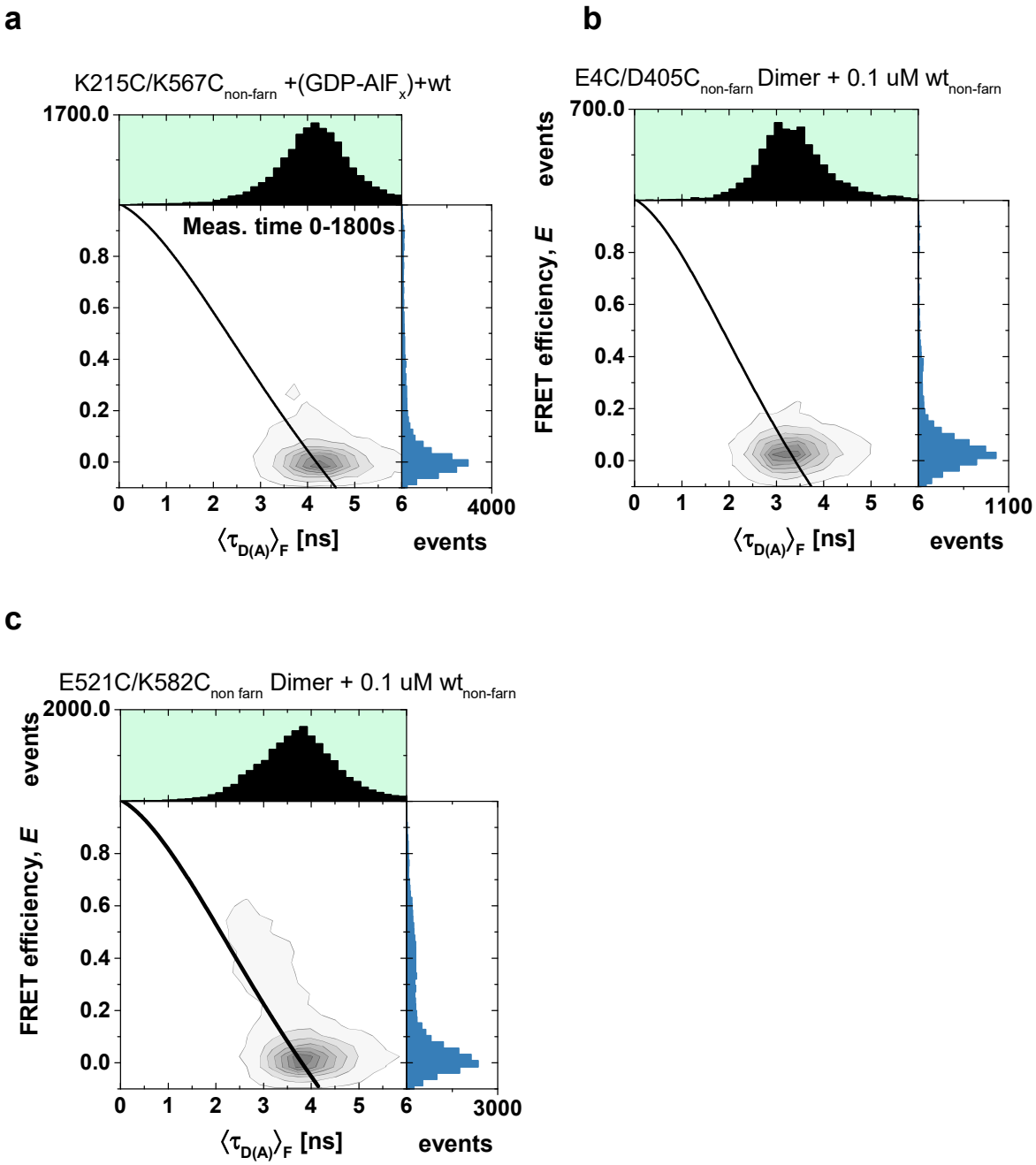
Mutant	R_{DA} [Å]	$R\pm$	6k1z	M_1	M_2
N18C/Q344C$_{farn}$	75	8	71	75	68
N18C/V577C$_{farn}$	58	5	55	55	42
E533C/M583C$_{farn}$	33	6	31	44	55
Q344C/A496C$_{farn}$	47	5	50	55	47
Q344C/Q525C$_{farn}$	48	5	39	40	36
χ^2_{red}			0.92	1.98	5.91

Supplementary Table 15: Overview of the distances in oligomeric state

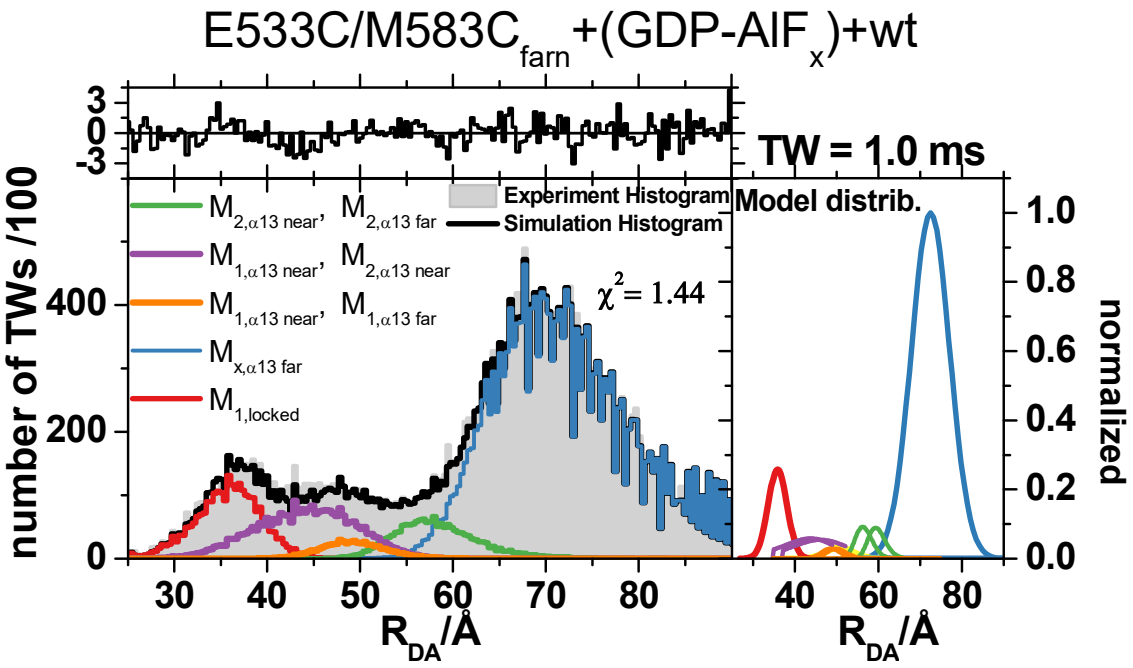
Mutant	seTCSPC	PDA	seTCSPC Imaging	Projection on N18C/V577C	6k1z α12/13 wide angle
	R_{DA} [Å]	R_{DA} [Å]	R_{DA} [Å]	R_{DA} [Å]	R_{DA} [Å]
N18C/Q344C$_{farn}$	69	69	71	65	66
N18C/V577C$_{farn}$			280 (from [29])	-	242
E533C/M583C$_{farn}$	85	83	80	77	73
Q344C/A496C$_{farn}$	75	69	72	66	71
Q344C/Q525C$_{farn}$	93	86	-	not used	108



Supplementary Figure 28: 2d-MFD-histograms of two farnesylated hGBP1 samples using different nucleotides. Top histogram shows the distribution of fluorescence weighted donor lifetimes in presence of an acceptor, $\langle\tau\rangle_F$. Right histogram shows the distribution of FRET efficiencies. Only bursts are shown from double labeled molecules, filtering was done using a stoichiometry and ALEX2CDE cut. Black line represents the static FRET-line, dashed magenta line the bleaching curve. a,b,c shows the N18C/Q577C$_{farn}$-sample adding no nucleotide, GMP and wild type and GDP-AlF$_x$, respectively. c shows the E533C/M583C$_{farn}$ sample with GTPγS and wild type.



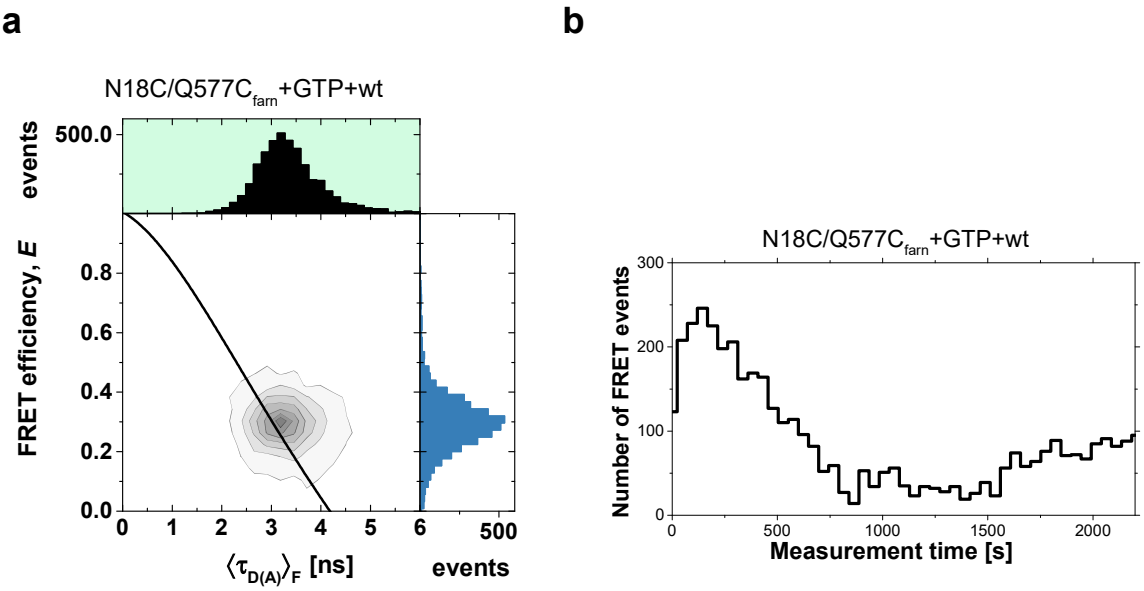
Supplementary Figure 29: Dimerization of non-farnesylated hGBP1. 2d-MFD of **a** K215C/K567C, **b** E4C/D405C and **c** E521C/K582C under the addition of GDP-AlF$_x$ and 0.1 µM wild type. Static FRET line is indicated in black.



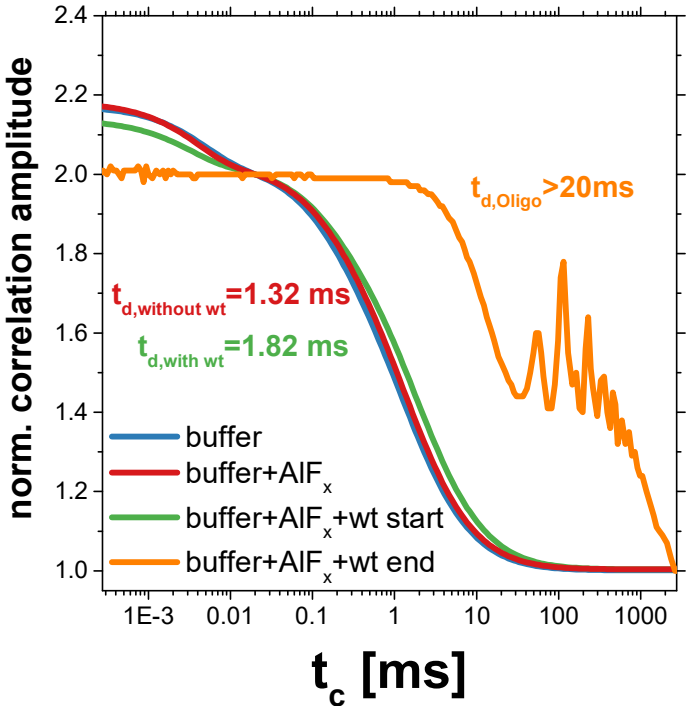
Supplementary Figure 30: PDA of E533C/M583C$_{farn}$ with nucleotide GDP-AlF$_x$ and additional wild type. Analysis was done globally over different TWs, here TW 1 ms is shown. Measured FRET derived distances R_{DA} were histogramed and fitted using the same model as for the non-farnesylated hGBP1. Model distribution is shown on the right-hand side of the panel

Supplementary Table 16: PDA of E533C/M583C$_{farn}$ with nucleotide GDP-AlF$_x$ and wild type.

	Fraction [%]					Relaxation times [µs]		
Sample	$M_{1,at}$	$M_{1,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ near}$	$M_{1,\alpha13\ far}$ ⇋ $M_{1,\alpha13\ near}$	$M_{2,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ far}$	$M_{x,\alpha13\ far}$	$M_{1,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ near}$	$M_{1,\alpha13\ far}$ ⇋ $M_{1,\alpha13\ near}$	$M_{2,\alpha13\ near}$ ⇋ $M_{2,\alpha13\ far}$
E533C/M583C+GDP-AlF$_x$+wt	9.1	11.8	2.4	7.1	69.6	270	2	3

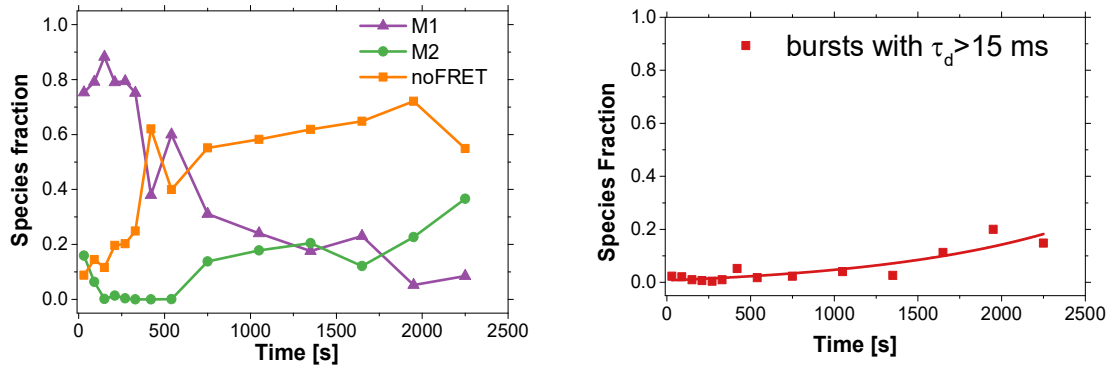


Supplementary Figure 31: a 2d-MFD-histograms of N18C/Q577C$_{farn}$ with GTP as a nucleotide and additional wt. **b** Number of FRET events with a FRET efficiency greater than zero as a function of the measurement time.



Supplementary Figure 32: FCS of farnesylated hGBP1 samples with nucleotide GDP-AlF_x and additional wild type.

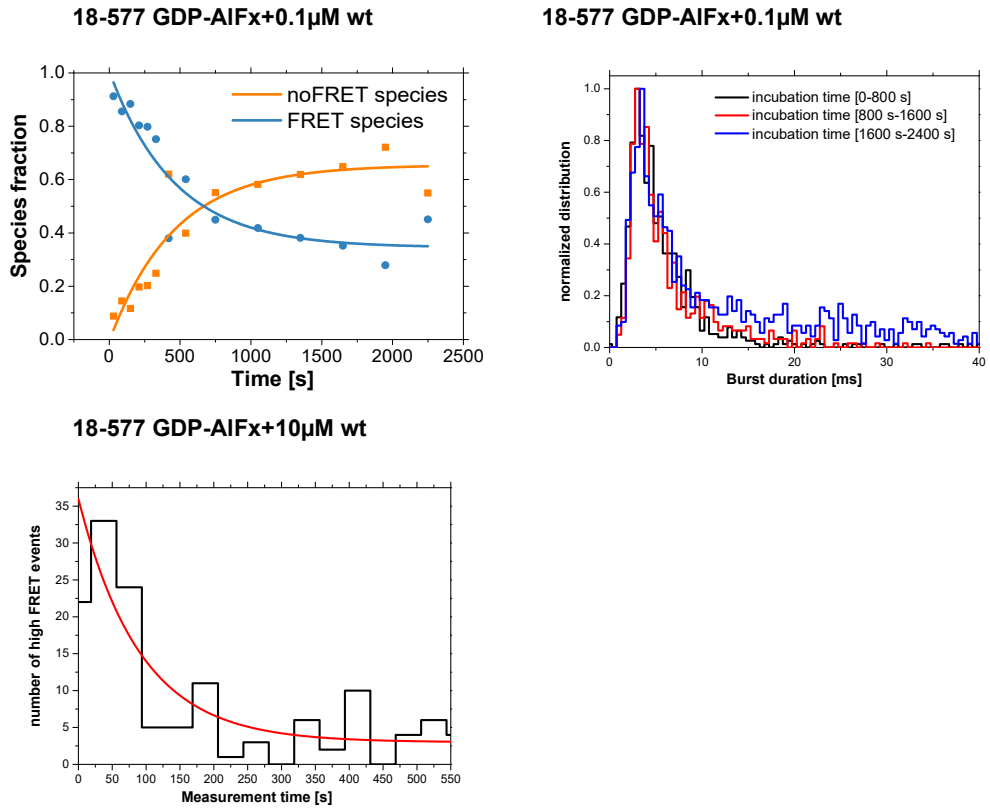
a **b**



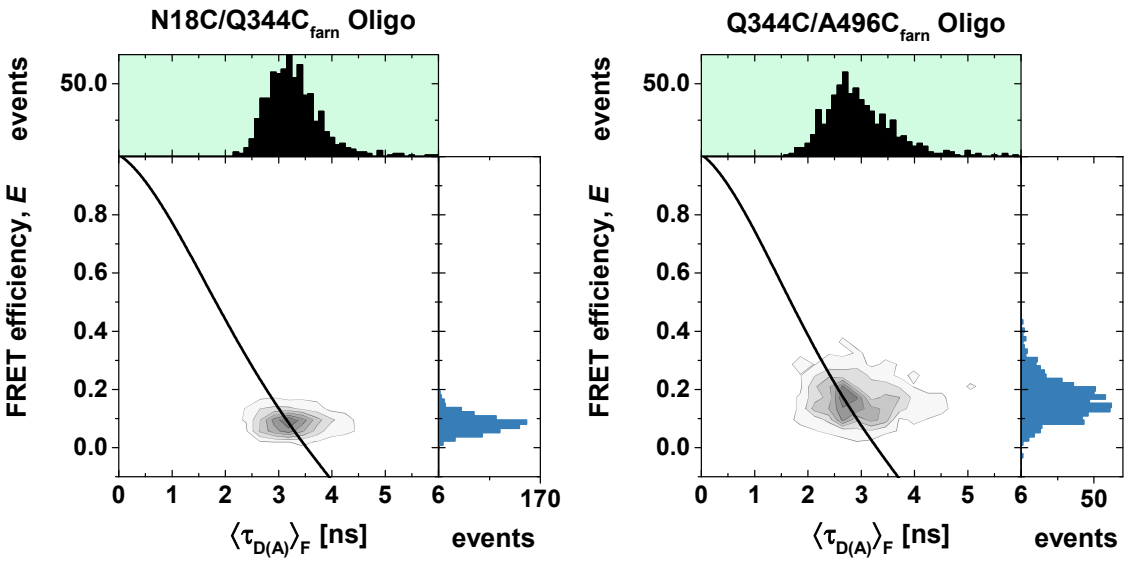
Supplementary Figure 33: N18C/Q577C$_{farn}$ seTCSPC and burst duration analysis under addition GDP-AlF$_x$ and wild type. a seTCSPC analysis where for each time interval 3 Gausian distributions were used. Purple dots are the species fractions of the M_1-state, green of the M_2-state and orange of the noFRET species. b Ratio of bursts with a burst duration higher than 15 ms in ratio to the total amount of bursts in dependence of the measurement time.

Supplementary Table 17: Results of seTCSPC fit of N18C/Q577C$_{farn}$ under addition GDP-AlF$_x$ and wild type.

Measurement Interval [s]	$R_{DA,1}$ [Å]	x_1	$R_{DA,2}$ [Å]	x_2	$R_{DA,3}$ [Å]	x_3	ratio of mol. with τ_d>15 ms
0-60	65	0.75	46	0.16	111	0.09	0.02
60-120	65	0.79	46	0.06	111	0.14	0.02
120-180	65	0.88	46	0.00	111	0.12	0.01
180-240	65	0.79	46	0.01	111	0.20	0.01
240-300	65	0.79	46	0.00	111	0.20	0.00
300-360	65	0.75	46	0.00	111	0.25	0.01
360-480	65	0.38	46	0.00	111	0.62	0.05
480-600	65	0.60	46	0.00	111	0.40	0.02
600-900	65	0.31	46	0.14	111	0.55	0.02
900-1200	65	0.24	46	0.18	111	0.58	0.04
1200-1500	65	0.18	46	0.20	111	0.62	0.03
1500-1800	65	0.23	46	0.12	111	0.65	0.11
1800-2100	65	0.05	46	0.23	111	0.72	0.20
2100-2400	65	0.08	46	0.37	111	0.55	0.15

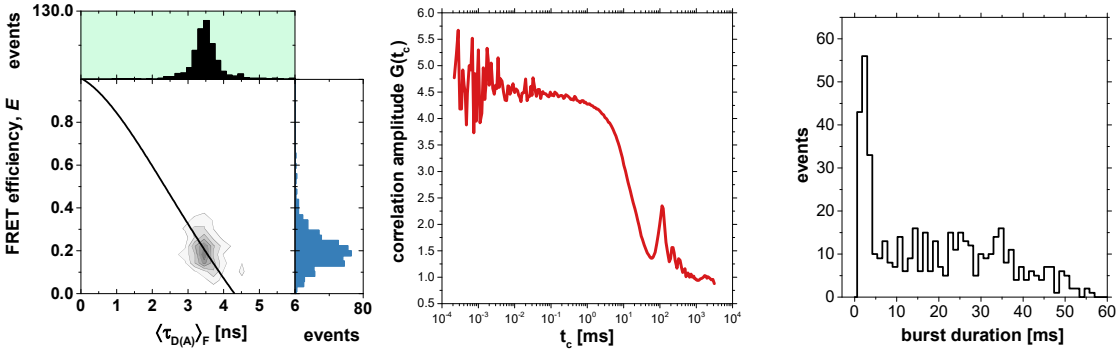


Supplementary Figure 34: Dimerization of N18C/Q577C farnesylated hGBP1. a Time resolved dimerization fitted with an exponential kinetic with one lifetime. Fraction of no- and FRET species was estimated with a bin time of 1 minute (see **Supplementary Table 17**). **b** Distribution of burst durations for the beginning of the measurement (black), in the middle of the measurement (red) and at the end of the measurement (blue). **c** Histogram of the number of high FRET events fitted with an exponential decay.

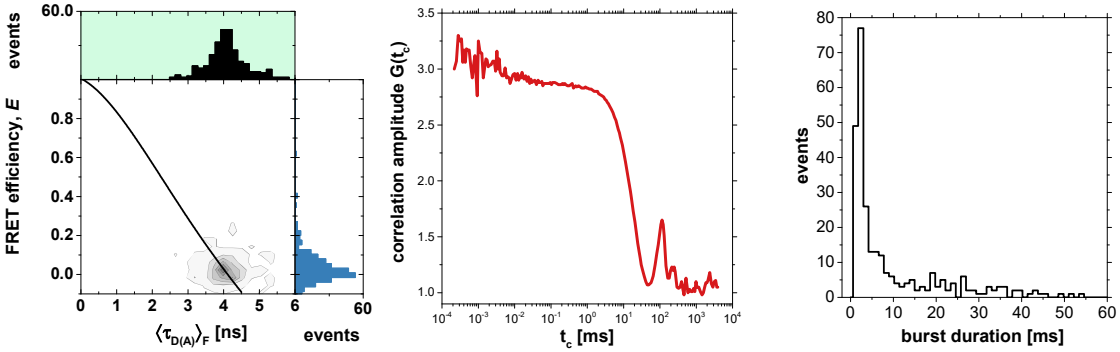


Supplementary Figure 35: Oligo-measurement of hGBP1 using fluorescence imaging. 2D sm-MFD histogram of the measurement with the distribution of lifetimes of the donor in presence of an acceptor $\langle\tau_{D(A)}\rangle_f$ (top, black histogram) and the distribution of FRET efficiency values *E* (right, blue distribution). Black line is representing the static FRET line.

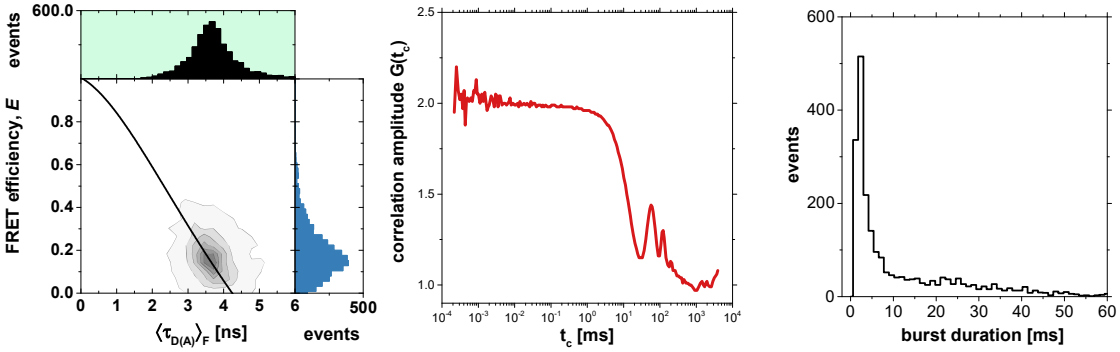
N18C/Q344C



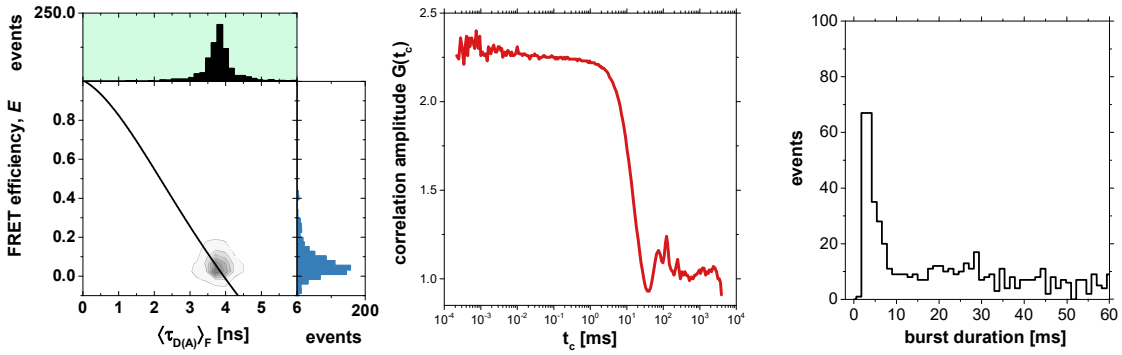
N18C/Q577C



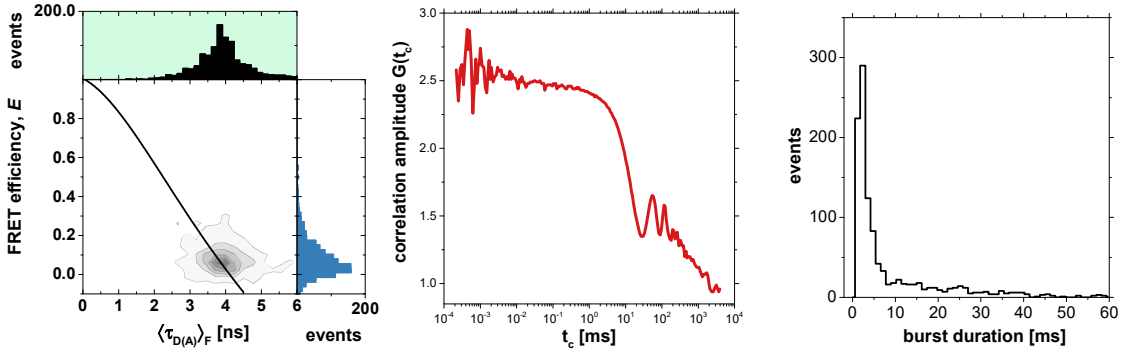
Q344C/A496C



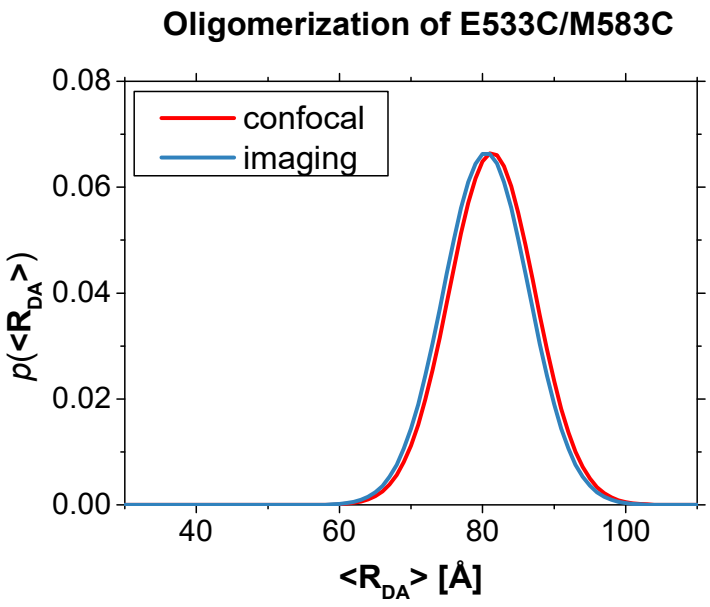
Q344C/Q525C



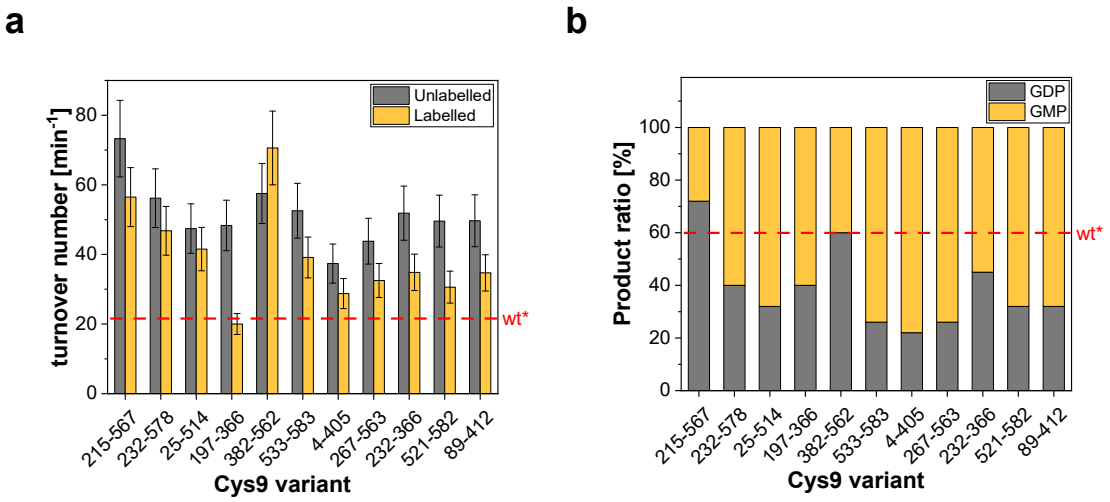
E533C/M583C



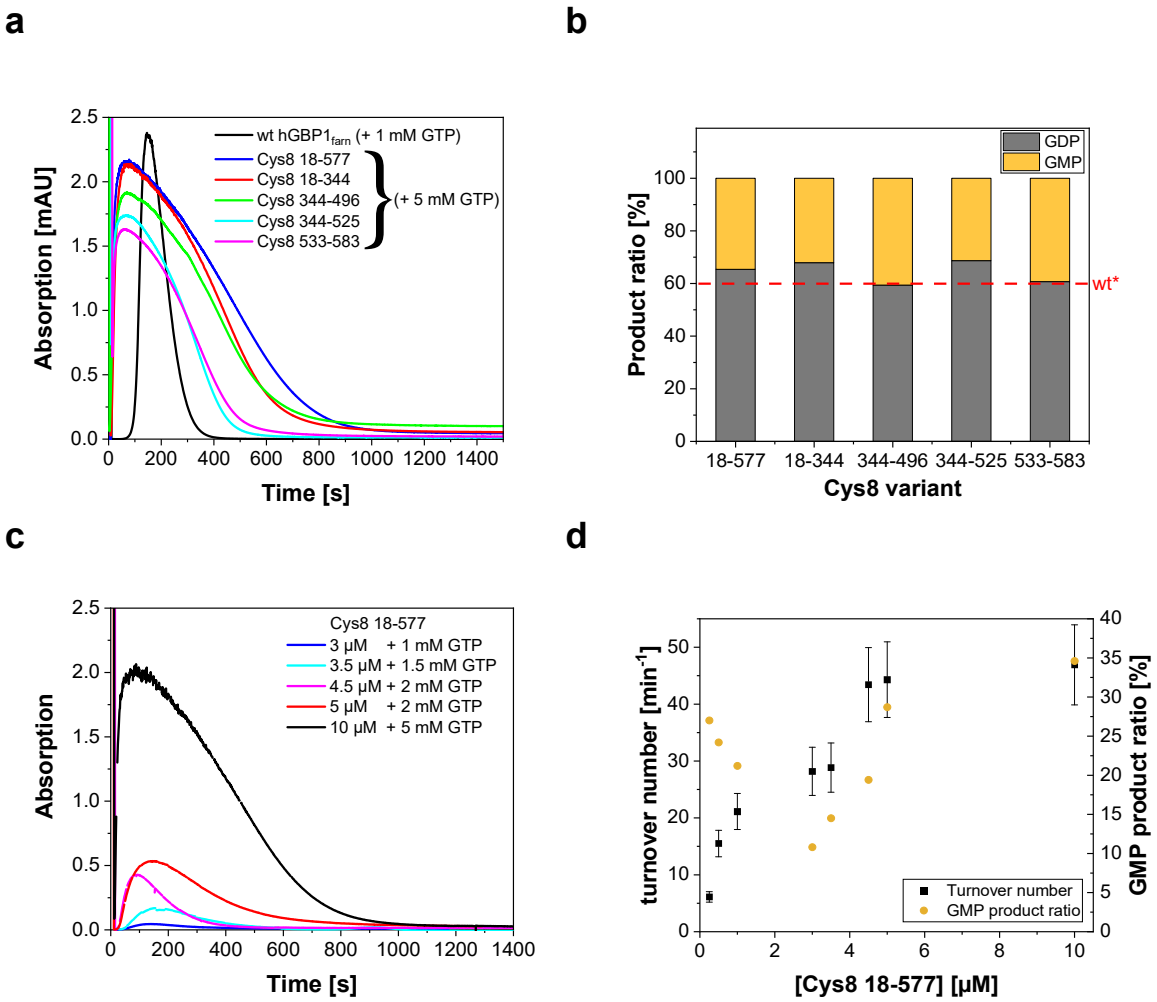
Supplementary Figure 36: Oligo-measurement of hGBP1 using confocal fluorescence spectroscopy. Farnesylated hGBP1 with nucleotide (GDP-AlF$_x$) and a high amount (10µM) of unlabeled farnesylated wt-hGBP1. Each panel shows a 2d-histogram for the particular sample (left), an FCS analysis (middle) and a distribution of burst durations (right).



Supplementary Figure 37: Comparison confocal to imaging. Distance distribution using a Gaussian model in a seTCSPC analysis for confocal measurement (red) and imaging (blue) of the sample E533C/M583C.



Supplementary Figure 38: Hydrolysis data of measured Cys9 samples. a Turnover number of Cys9 variants, in both the unlabeled (gray) and labeled (yellow) state, hydrolysing GTP at 25 °C. Each measurement involved 1 µM of protein with 0.5 mM of GTP in a buffer of 50 mM Tris(HCl), 150 mM NaCl and 5 mM $MgCl_2$ at pH 7.9. Samples were taken and stopped at specified time intervals to analyze the ratios of GTP, GDP and GMP within the sample at the given time. As the hydrolysis speed was saturated due to the excess of GTP in solution, the decline of GTP could be linearly fitted. Using the slope of this linear fit, the turnover number could be calculated using $turnover\ number\ [min^{-1}] = \frac{\frac{d(GTP\%)}{dt} \cdot [GTP]_0}{[hGBP1] \cdot 100} \cdot 60$. The red line symbolizes the turnover number 21.6 min^{-1} of wild type hGBP1 at saturated dimerization conditions [40]. The given error for each measurement is 15% of the acquired turnover number, as determined by previous research [18, 41] . **b** Final product ratios of GDP (gray) and GMP (yellow) following hydrolysis experiments depicted in a. The red line shows the product ratio of wild type hGBP1 of 60% GDP and 40% GMP [40].



Supplementary Figure 39: GTP hydrolysis of Cys8 variants. a Turbidity assays of wild type hGBP1 (black) and Cys8 variants (colored). For all experiments 10 µM of protein in buffer containing 50 mM Tris-HCl, 150 mM NaCl and 5 mM MgCl2 at pH 7.9 was used. As the hydrolysis rate was faster for Cys8 variants, it was compensated by adding 5 mM of GTP instead of the 1 mM GTP for wild type. The turbidity was measured as absorbance at 350 nm. b Final product ratios of GDP (gray) and GMP (yellow) after turbidity assays shown in a. The product ratio of wild type is shown as the red dotted line with 60% GDP and 40% GMP. c Turbidity assays of variant Cys8 18-577 at different concentrations and varying concentrations of GTP. d Maximum turnover numbers of hydrolysis reactions with Cys8 18-577 acquired the same way as turnover numbers shown in **Supplementary Figure 38**a using $\boldsymbol{turnover\ number}\ [\boldsymbol{min^{-1}}] = \frac{\frac{\boldsymbol{d(GTP\%)}}{\boldsymbol{dt}} \cdot [\boldsymbol{GTP}]_{\boldsymbol{0}}}{[\boldsymbol{hGBP1}] \cdot \boldsymbol{100}} \cdot \boldsymbol{60}$.. All reactions which also displayed turbidity are shown in c (lowest concentration with turbidity being 3 µM of protein). Additionally, the final product ratio of GMP is shown as a yellow dot for each concentration.

Supplementary notes

Notes regarding the turbidity of Cys8 variants:

Turbidity experiments of Cys8 variants showed that no lagtime occurs while polymerizing GTP, which is different from wild type hGBP1 (**Supplementary Figure 39**). Other hGBP1 variants which exhibited no lag time in polymerization in literature were proven to have a weakened interaction between the LG and GED domains (e. g. variant R227E/K228E [32]). This however, is not the case here, as was shown that the sole addition of GDP-AlF_x does not lead to a separation of the GED from the rest of the protein (see **Figure 43d**). Additionally, the final product ratio of GDP/GMP is very similar to that of wild type hGBP1 (see **Supplementary Figure 39**). Therefore, it was concluded that the lag of a lagtime before polymerization stems from the heightened hydrolysis rate by removing natural cysteines within the protein.

Instrument used for confocal microscopy:

Experiments were performed on a custom-designed Abberior Instruments Expert Line microscope (Abberior Instruments, Göttingen, Germany) with the following components: An Olympus IX83 microscope body equipped with an *easy 3D* module based on a spatial light modulator (SLM). Excitation lasers with wavelengths of 488 nm and 640 nm with a pulse width of <100 ps were used to excite Alexa 488 and Alexa 647 respectively. Photon arrival was time resolved using an external TCSPC unit (Hydra Harp 400, PicoQuant GmbH, Germany). All confocal imaging FRET experiments were performed with a water-immersion objective (UPlanSApo 60x/1.2w, Olympus, Germany).

Instrument used for confocal fluorescence spectroscopy

For sm-MFD measurement an inverted microscope (Olympus IX70) was used. As an objective we used the Olympus UPlanSAp 60x/1.2. For excitation a linearly polarized pulsed diode laser was used, using for green excitation (λ=485 nm) LDH-D-C 485, PicoQuant, and for red excitation (λ=640 nm) LDH-D-C 640, Picoquant. Both were operated in PIE-configuration with each 32 MHz. Fluorescence signal was filtered from excitation light using a triple band beamsplitter 488/570/640 (AHF Analysentechnik). A polarized beam splitter was used to split detected light into perpendicular and parallel polarization. Color filtering was done using ET 535/50 and HQ 730/140 (AHF Analysentechnik). Signal was detected with 8 APD detectors. Single photon counting was done using synchronized channels (HydraHarp 400, PicoQuant) operating in Time-Tagged Time-Resolved (TTTR) mode. Data analysis was done using in-house software available upon request on the website of the Seidel group (https://www.mpc.hhu.de/software/mfd-fcs-and-mfis).